

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/006473

International filing date: 01 March 2005 (01.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/548,191
Filing date: 01 March 2004 (01.03.2004)

Date of receipt at the International Bureau: 18 April 2005 (18.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1300405

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 23, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/548,191

FILING DATE: *March 01, 2004*

RELATED PCT APPLICATION NUMBER: *PCT/US05/06473*



Certified by

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office



13281 U.S. PTO
030104

PROVISIONAL APPLICATION COVER SHEET

Mail Stop Provisional Patent Application

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(c).

Docket Number 8940.6144		Type a plus sign (+) inside this box →	
INVENTOR(s)/APPLICANT(s)			
LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Lee Hestir Chu Williams	Ernestine Kevin Keting Lewis		Castro Valley, CA Castro Valley, CA Burlingame, CA Mill Valley, CA
TITLE OF INVENTION (500 characters max)			
NOVEL HUMAN cDNA CLONES AND METHODS OF THEIR USE			
CORRESPONDENCE ADDRESS			
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P. 1300 I Street, N.W. Washington, D.C. 20005 Telephone No. (202) 408-4000 Customer Number 22,852			
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification	461 Pages	<input type="checkbox"/> Small Entity Statement	
<input type="checkbox"/> Drawing(s)	____ Sheets ____ Figures	<input type="checkbox"/> Other (specify)	
METHOD OF PAYMENT (check one)			
<input type="checkbox"/> A check or money order is enclosed to cover the Provisional filing fees		PROVISIONAL FILING FEE	
<input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number 06-0916.		<input checked="" type="checkbox"/> \$160.00 <input type="checkbox"/> \$80.00 (small entity)	

22151 U.S. PTO
60/548191



030104

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are:

Respectfully submitted,

SIGNATURE

Jean B. Fordis
Reg. No. 53,283

Date March 1, 2004

TYPED OR PRINTED NAME Jean B. Fordis

REGISTRATION NO. 32,984

☐ Additional inventors are being named on separately numbered sheets attached hereto.

PROVISIONAL APPLICATION FILING ONLY

UNITED STATES PROVISIONAL PATENT APPLICATION

for

NOVEL HUMAN cDNA CLONES AND METHODS OF
THEIR USE

by

ERNESTINE LEE

KEVIN HESTIR

KETING CHU

and

LEWIS T. WILLIAMS

BACKGROUND OF THE INVENTION**Field of the Invention**

[001] The present invention is related generally to novel human cDNA clones and novel polypeptides encoded therefrom, and their compositions. The polynucleotides and polypeptides are useful in diagnostic and therapeutic applications for a variety of diseases and conditions. The present invention also relates to methods of modulating a biological activity through the use of the novel polynucleotides and novel polypeptides of the invention.

Description of Related Art

[002] Sequencing of the genomes, or portions of the genomes, of numerous biological materials, including humans, animals, microorganisms and various plant varieties, has been, and is being carried out on a large scale. Polynucleotides identified using sequencing techniques may be partial or full-length genes, and may contain open reading frames, or portions of open reading frames, that encode polypeptides. Putative polypeptides may be determined based on polynucleotide sequences. The sequencing data relating to polynucleotides thus represents valuable and useful information.

[003] Polynucleotides may be analyzed for various degrees of novelty by comparing identified sequences to sequences published in various public domain databases, such as EMBL. Newly identified polynucleotides and putative polypeptides may also be compared to polynucleotides and polypeptides contained in public domain information to ascertain homology to known polynucleotides and polypeptides. In this way, the degree of similarity, identity, or homology of polynucleotides and polypeptides of unknown function may be determined relative to polynucleotides and polypeptides having known functions.

[004] Information relating to the sequences of isolated polynucleotides may be used in a variety of ways. Specified polynucleotides having a particular sequence may be isolated, or synthesized, for use in *in vivo* or *in vitro* experimentation as probes or primers. Alternatively, collections of sequences of isolated polynucleotides may be stored using magnetic or optical storage medium, and analyzed or manipulated using computer hardware and software, as well as other types of tools.

TECHNICAL FIELD

[005] The present invention is related generally to novel polynucleotides and novel polypeptides encoded thereby, their compositions, antibodies directed thereto, and other agonists or antagonists thereto. The polynucleotides and polypeptides are useful in

diagnostic, prophylactic, and therapeutic applications for a variety of diseases, disorders, syndromes, and conditions, as well as in discovering new diagnostics, prophylactics, and therapeutics for such diseases, disorders, syndromes, and conditions (hereinafter disorders). The present invention also relates to methods of modulating biological activities through the use of the novel polynucleotides and novel polypeptides of the invention and through the use of agonists and antagonists, such as antibodies, thereto.

[006] This application further relates to the field of polypeptides that are associated with regulating cell growth and differentiation, that are over-expressed in cancer, and/or that can be associated with proliferation or inhibition of cancer growth, including hematopoietic cancers such as leukemias, lymphomas, and solid cancers such as pancreatic cancer, tracheal cancer, and lung cancer, for example, adenocarcinomas and/or squamous cell carcinomas. These polypeptides may also be associated with other conditions, such as inflammatory, immune, and metabolic disorders, as well as microbial infections, including viral, bacterial, fungal, and parasitic disorders.

[007] This application further relates to modulators of biological activity that can specifically bind to these polynucleotides or polypeptides, or otherwise specifically modulate their activity. For example, they can directly or indirectly induce antibody-dependent cellular cytotoxicity (ADCC), complement-dependent cytotoxicity (CDC), endocytosis, apoptosis, or recruitment of other cells to effect cell activation, cell inactivation, cell growth or differentiation or inhibition thereof, and cell killing.

SUMMARY OF THE INVENTION

1. A first nucleic acid molecule comprising a polynucleotide sequence chosen from at least one polynucleotide sequence according to SEQ. ID. NOS. 1-187; SEQ. ID. NOS. 375-484, or a complement thereof, or from at least one polypeptide sequence that encodes SEQ. ID. NOS. 188-374.
2. The nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA or a RNA molecule.
3. An animal injected with the nucleic acid molecule of claim 1.
4. A double-stranded isolated nucleic acid molecule comprising the first nucleic acid molecule of claim 1 and its complement.
5. The nucleic acid molecule of claim 4, wherein the first polynucleotide sequence encodes a polypeptide chosen from a polypeptide comprising a signal peptide, a

mature polypeptide that lacks a signal peptide, a signal peptide, a biologically active fragment of a polypeptide, a polypeptide lacking a signal peptide cleavage site, a polypeptide consisting essentially of a N-terminal fragment that contains a Pfam domain, and a polypeptide consisting essentially of a C-terminal fragment that contains a Pfam domain.

6. A second nucleic acid molecule comprising a second polynucleotide sequence that is at least about 70%, or about 80%, or about 90%, or about 95% homologous to the first nucleic acid molecule of claim 1.

7. A second isolated nucleic acid molecule comprising a second polynucleotide sequence that hybridizes to the first polynucleotide sequence of claim 1 under high stringency conditions.

8. The second isolated nucleic acid molecule of claim 6, wherein the second polynucleotide sequence is complementary to the first polynucleotide sequence.

9. A vector comprising the nucleic acid molecule of claim 1 and a promoter that drives the expression of the nucleic acid molecule.

10. The vector of claim 9, wherein the promoter is chosen from one or more of a promoter that is naturally contiguous to the nucleic acid molecule, a promoter that is not naturally contiguous to the nucleic acid molecule, an inducible promoter, a conditionally active promoter, a constitutive promoter, and a tissue specific promoter.

11. A host cell transformed, transfected, transduced, or infected with the nucleic acid molecule of claim 1.

12. The host cell of claim 11, wherein the cell is chosen from one or more of a prokaryotic cell, a eucaryotic cell, a human cell, a mammalian cell, an insect cell, a fish cell, a plant cell, and a fungal cell.

13. A nucleic acid composition comprising a pharmaceutically acceptable carrier or a buffer and one or more compositions chosen from the nucleic acid molecule of claim 1, the nucleic acid molecule of claim 4, the vector of claim 9, and the host cell of claim 11.

14. A substantially purified polypeptide comprising a polypeptide sequence chosen from at least one amino acid sequence according to SEQ. ID. NOS. 188-374.

15. An animal injected with the polypeptide of claim 14.

16. The polypeptide of claim 14, wherein the polypeptide has a function chosen from an agonist, an antagonist, a ligand, and a receptor.

17. The polypeptide of claim 14, wherein the polypeptide is chosen from a polypeptide comprising a signal peptide, a mature polypeptide that lacks a signal peptide, a

signal peptide, a biologically active fragment of a polypeptide, a polypeptide lacking a signal peptide cleavage site, a biologically active fragment consisting essentially of an N-terminal fragment containing a Pfam domain, a biologically active fragment consisting essentially of a C-terminal fragment containing a Pfam domain, an extracellular fragment, a ligand binding fragment, and a receptor binding fragment.

18. A polypeptide composition comprising the polypeptide molecule of claim 14 and a pharmaceutically acceptable carrier or a buffer.

19. A cell culture medium comprising the polypeptide of claim 14.

20. The cell culture medium of claim 19, further comprising responder cells chosen from one or more T cells, B cells, NK cells, dendritic cells, macrophages, muscle cells, stem cells, epithelial skin cells, fat cells, blood cells, brain cells, bone marrow cells, endothelial cells, retinal cells, bone cells, kidney cells, pancreatic cells, liver cells, spleen cells, prostate cells, cervical cells, ovarian cells, breast cells, tracheal cells, lung cells, liver cells, soft tissue cells, colorectal cells, cells of the gastrointestinal tract, and cancer cells.

21. The cell culture medium of claim 20, wherein the responder cells proliferate in the medium.

22. The cell culture medium of claim 20, wherein the responder cells are inhibited in the medium.

23. A cell culture comprising transfected cells, wherein the transfected cells are transfected with the polynucleotide of claim 1.

24. The cell culture of claim 23, further comprising responder cells chosen from one or more T cells, B cells, NK cells, dendritic cells, macrophages, muscle cells, stem cells, epithelial skin cells, fat cells, blood cells, brain cells, bone marrow cells, endothelial cells, retinal cells, bone cells, kidney cells, pancreatic cells, liver cells, spleen cells, prostate cells, cervical cells, ovarian cells, breast cells, tracheal cells, lung cells, liver cells, soft tissue cells, colorectal cells, cells of the gastrointestinal tract, and cancer cells.

25. The cell culture of claim 23, wherein the responder cells proliferate in the cell culture.

26. The cell culture of claim 23, wherein the responder cells are inhibited in the cell culture.

27. A method of making a transformed, transfected, transduced, or infected host cell comprising:

(a) providing a composition comprising the vector of claim 9, and

(b) allowing a host cell to come into contact with the vector to form a transformed, transfected, transduced, or infected host cell.

28. A method of making a polypeptide comprising:

- (a) providing a nucleic acid molecule that comprises a polynucleotide sequence encoding the polypeptide of claim 14;
- (b) introducing the nucleic acid molecule into an expression system; and
- (c) allowing the polypeptide to be produced.

29. A method of making a polypeptide comprising:

- (a) providing a composition comprising the host cell of claim 11;
- (b) culturing the host cell to produce the polypeptide; and
- (c) allowing the polypeptide to be produced.

30. A diagnostic kit comprising a polynucleotide molecule, wherein the polynucleotide molecule comprises a sequence chosen from (a) at least 6, (b) at least 7, (c) at least 8, and (d) at least 9 contiguous nucleotides chosen from the nucleic acid molecule of claim 1.

31. A diagnostic kit comprising a polypeptide molecule, wherein the polypeptide molecule comprises an amino acid sequence or a biologically active fragment thereof, derived from the nucleic acid molecule of claim 1.

32. A genetically modified mouse comprising a deletion, substitution, or modification of a sequence chosen from SEQ. ID. NOS. 1-187; SEQ. ID. NOS. 375-484, wherein the deletion, substitution or modification prevents or reduces expression of said sequence and results in a mouse deficient in or completely lacking one or more gene products of a sequence chosen from SEQ. ID. NOS. 1-187; SEQ. ID. NOS. 375-484.

33. A method of determining the presence of the nucleic acid molecule of claim 1 or its complement comprising:

- (a) providing a complement to the nucleic acid molecule or providing a complement to the complement of the nucleic acid molecule;
- (b) allowing the molecules to interact; and
- (c) determining whether interaction has occurred.

34. A method of determining the presence of an antibody to the polypeptide of claim 14 in a sample, comprising:

- (a) providing the polypeptide;

(b) allowing the polypeptide to interact with any specific antibody in the sample;
and

(c) determining whether interaction has occurred.

35. An antibody specifically recognizing, binding to, and/or modulating the biological activity of at least one polypeptide encoded by a nucleic acid molecule of claim 1, or a biologically active fragment thereof.

36. An antibody composition comprising the antibody of claim 35 and a pharmaceutically acceptable carrier.

37. The antibody of claim 35, wherein the antibody is chosen from one or more of a monoclonal antibody, a polyclonal antibody, a single chain antibody, an antibody comprising a backbone of a molecule with an Ig domain, a targeting antibody, a neutralizing antibody, a stabilizing antibody, an enhancing antibody, an antibody agonist, an antibody antagonist, an antibody that promotes endocytosis of a target antigen, a cytotoxic antibody, an antibody that mediates ADCC, a human antibody, a non-human primate antibody, a non-primate animal antibody, a rabbit antibody, a mouse antibody, a rat antibody, a sheep antibody, a goat antibody, a horse antibody, a porcine antibody, a cow antibody, a chicken antibody, a humanized antibody, a primatized antibody, and a chimeric antibody.

38. The antibody of claim 37, wherein the antibody is produced in a manner chosen from *in vivo* and *in vitro*.

39. The antibody of claim 37, wherein the antibody is produced in an organism chosen from a prokaryote and a eukaryote.

40. The antibody of claim 39, wherein the organism is chosen from a bacterial cell, a fungal cell, a plant cell, an insect cell, and a mammalian cell.

41. The antibody of claim 40, wherein the cell is chosen from a yeast cell, an *Aspergillus* cell, an SF9 cell, a High Five cell, a cereal plant cell, a tobacco cell, and a tomato cell.

42. The cytotoxic antibody of claim 37, further comprising one or more cytotoxic component chosen from a radioisotope, a microbial toxin, a plant toxin, and a chemical compound.

43. The cytotoxic antibody of claim 42, wherein the chemical compound is chosen from doxorubicin and cisplatin.

44. The antibody of claim 35, wherein the antibody has a function chosen from specifically inhibiting the binding of the polypeptide to a ligand, specifically inhibiting the

binding of the polypeptide to a substrate, specifically inhibiting the binding of the polypeptide as a ligand, and specifically inhibiting the binding of the polypeptide as a substrate.

45. A bacteriophage, wherein the antibody of claim 35, or a fragment thereof, is displayed on the bacteriophage.

46. A bacterial cell comprising the bacteriophage of claim 45.

47. A non-human animal injected with the antibody composition of claim 36.

48. A host cell that secretes the antibody of claim 35.

49. A method of making an antibody, comprising:

(a) introducing a polypeptide, polynucleotide encoding the polypeptide, or a biologically active fragment thereof into an animal in sufficient amount to elicit generation of antibodies specific to the polypeptide, wherein the polypeptide:

(i) is encoded by the nucleic acid molecule of claim 1; or

(ii) comprises the polypeptide sequence of claim 14; and

(b) recovering the antibodies therefrom.

50. The method of claim 49, further comprising after step (a), the step of isolating a spleen from the animal injected with the polypeptide or polynucleotide or a fragment thereof, and the step of recovering the antibodies from the spleen cells.

51. The method of claim 50, further comprising the step of making a hybridoma using cells from the spleen and selecting a hybridoma that secretes the antibodies.

52. The method of claim 50, further comprising making a polynucleotide library from the spleen cells, selecting a cDNA clone that produces the antibodies, and expressing the cDNA clone in an expression system to produce antibodies or fragments thereof.

53. A method of modulating biological activity comprising:

(a) providing the antibody of claim 35; and

(b) contacting the antibody with a first human or a non-human host cell thereby modulating the activity of a first human or non-human animal host cell, or a second host cell.

54. The method of claim 53, wherein the modulation of biological activity is chosen from enhancing cell activity directly, enhancing cell activity indirectly, inhibiting cell activity directly, and inhibiting cell activity indirectly.

55. The method of claim 53, wherein the step of contacting the antibody with a first human or non-human host cells results in recruitment of the second host cell.

56. The method of claim 53, wherein the first host cell is a cancer cell.

57. The method of claim 53, wherein the first or second host cell is chosen from a T cell, B cell, NK cell, dendritic cell, macrophage, muscle cell, stem cell, skin cell, fat cell, blood cell, brain cell, bone marrow cell, endothelial cell, retinal cell, bone cell, kidney cell, pancreatic cell, liver cell, spleen cell, prostate cell, cervical cell, ovarian cell, breast cell, tracheal cells, lung cell, liver cell, soft tissue cell, colorectal cell, and gastrointestinal tract cell.

58. A method of diagnosing a disease, disorder, syndrome, or condition chosen from cancer, proliferative, inflammatory, immune, metabolic, genetic, bacterial, and viral diseases, disorders, syndromes, or conditions in a patient, comprising:

- (a) providing the antibody of claim 35;
- (b) allowing the antibody to contact a patient sample; and
- (c) detecting specific binding between the antibody and an antigen in the sample to determine whether the subject has cancer, a proliferative, inflammatory, immune, metabolic, genetic, bacterial, or viral disease, disorder, syndrome, or condition.

59. A method of diagnosing a disease, disorder, syndrome, or condition chosen from cancer, proliferative, inflammatory, immune, bacterial, and viral diseases, disorders, syndromes, or conditions in a patient, comprising:

- (a) providing a polypeptide that specifically binds the antibody of claim 35;
- (b) allowing the polypeptide to contact a patient sample; and
- (c) detecting specific binding between the polypeptide and any interacting molecule in the sample to determine whether the subject has cancer, a proliferative, inflammatory, immune, bacterial, or viral disease, disorder, syndrome, or condition.

60. A method of identifying an agent that modulates the biological activity of a polypeptide comprising:

- (a) providing a polypeptide or an active fragment thereof, wherein the polypeptide comprises at least one amino acid sequence according to SEQ. ID. NOS. 188-374;
- (b) allowing at least one agent to contact the polypeptide; and
- (c) selecting an agent that binds the polypeptide or affects the biological activity of the polypeptide.

61. The method of claim 60, wherein the polypeptide is expressed on a cell surface.
62. A modulator composition comprising a modulator and a pharmaceutically acceptable carrier, wherein the modulator is obtainable by the method of claim 60.
63. The modulator composition of claim 62, wherein the modulator is an antibody.
64. A method of treating a disease, disorder, syndrome, or condition in a subject, comprising administering the composition of any one of claims 13, 18, and 36 to the subject.
65. The method of claim 64, wherein the composition is administered in a manner chosen from orally, parenterally, by implantation, by inhalation, intranasally, intravenously, intra-arterially, intracardiacally, subcutaneously, intraperitoneally, transdermally, intraventricularly, intracranially, and intrathecally.
66. The method of claim 64, wherein the disease, disorder, syndrome, or condition is chosen from cancer, a proliferative, inflammatory, immune, metabolic, genetic, bacterial, and viral disease, disorder, syndrome, or condition.
67. The method of claim 64, wherein the disease is cancer.
68. A method of treating a disease, disorder, syndrome, or condition chosen from cancer, proliferative, inflammatory, immune, metabolic, genetic, bacterial, and viral diseases, disorders, syndromes, or conditions in a subject, comprising:
- (a) providing an antibody composition that comprises a first antibody or fragment thereof that specifically binds to a first epitope of a first polypeptide or a biologically active fragment thereof, wherein the first polypeptide:
 - (i) is encoded by the nucleic acid molecule of claim 1; or
 - (ii) comprises the polypeptide of claim 14; and
 - (b) administering the antibody composition to the subject.
69. The method of claim 68, wherein the antibody composition further comprises a second antibody that binds specifically to or interferes with the activity of a second epitope of the first polypeptide or to a first epitope of a second polypeptide.
70. The method of claim 69, wherein the second polypeptide comprises the polypeptide of 14.
71. A kit comprising the antibody of claim 35 and instructions for its use.
72. A method of gene therapy, comprising:
- (a) providing a polynucleotide comprising a nucleic acid molecule encoding the antibody of claim 35; and

(b) administering the polynucleotide to a subject.

73. A method for prophylactic or therapeutic treatment of a subject, comprising:

(a) providing a vaccine; and

(b) administering the vaccine to the subject;

wherein the vaccine comprises a polynucleotide or a polypeptide chosen from at least one sequence according to SEQ. ID. NOS. 1 - 484 or a biologically active fragment thereof.

74. The method of claim 73, wherein the vaccine is a cancer vaccine, and the polypeptide is a cancer antigen.

75. A method of inhibiting transcription or translation of a first polynucleotide encoding a first polypeptide, comprising:

(a) providing a second polynucleotide that hybridizes to the first polynucleotide, wherein the first polynucleotide comprises a polynucleotide sequence chosen from:

(i) at least one polynucleotide sequence according to SEQ. ID. NOS. 1-187 or 375-484;

(ii) a polynucleotide encoding a polypeptide comprising an amino acid sequence chosen from at least one amino acid sequence according to SEQ. ID. NOS. 188-374; and

(iii) a polynucleotide encoding a fragment of a polypeptide comprising an amino acid sequence chosen from at least one amino acid sequence according to SEQ. ID. NOS. 188-374; and

(b) allowing the first polynucleotide to contact the second polynucleotide.

76. A method of treating a disease, disorder, syndrome, or condition comprising administering a modulator to a subject, wherein the modulator binds to a cell surface molecule that is over-expressed in the disease, disorder, syndrome, or condition, and is linked to the antibody of claim 35.

77. The method of claim 76, wherein the antibody is capable of initiating antibody-dependent cellular cytotoxicity.

78. The method of claim 76, wherein the disease, disorder, syndrome, or condition is cancer and the cell surface molecule is over-expressed in a cancer cell.

79. An isolated modified cell comprising at least one first heterologous nucleic acid molecule, wherein the first heterologous nucleic acid molecule comprises a first

polynucleotide sequence that encodes a first polypeptide, wherein the first polypeptide is chosen from the appendices and functions as a first therapeutic, diagnostic or prophylactic molecule for a disease, disorder, syndrome, or condition.

80. The cell of claim 79, wherein the disease, disorder, syndrome, or condition is chosen from cancer and other hyperproliferative diseases, CNS diseases, hematopoietic diseases, diabetes and other metabolic diseases, bone diseases, immune diseases, heart diseases, liver diseases, lung diseases, fertility-related diseases, kidney diseases, joint diseases, and related disorders, syndromes, or conditions.

81. The cell of claim 79, wherein the first polynucleotide sequence is chosen from a secreted molecule, an extracellular domain of a transmembrane molecule, a transmembrane molecule, and a receptor.

82. The cell of claim 79, wherein the first polynucleotide sequence is chosen from a growth factor, an anti-inflammatory factor, an anti-apoptotic factor, a vasodilator, an anti-coagulant, insulin, an anti-cancer agent, a signaling molecule, a biologically active fragment of the molecule, and a fusion protein comprising the molecule or biologically active fragment thereof.

83. The cell of claim 79, wherein the first polynucleotide sequence is chosen from an extracellular domain of a receptor, an immunoglobulin chain, a biologically active fragment of the molecule, and a fusion protein comprising the molecule or biologically active fragment thereof.

84. The cell of claim 79, wherein the first heterologous nucleic acid molecule encodes a fusion protein which is cleavable to release a biologically active molecule or biologically active fragment thereof.

85. The cell of claim 79, further comprising at least one second heterologous nucleic acid molecule that comprises at least one second polynucleotide sequence that encodes a second polypeptide.

86. The cell of claim 79, further comprising at least one second heterologous nucleic acid sequence encoding a second therapeutic molecule that functions in conjunction with the first therapeutic molecule.

87. The cell of claim 82, wherein the first therapeutic molecule is a growth factor chosen from a factor that stimulates cell proliferation, a factor that stimulates cell differentiation, a factor that promotes change in the metabolic state of a cell, and a factor that promotes change in the gene expression pattern of a cell.

88. The cell of claim 87, wherein the change in the metabolic state comprises activation, inactivation, or inhibition.
89. The cell of claim 79, wherein the first polypeptide is a growth factor.
90. The cell of claim 79, wherein the first polypeptide is an anti-inflammatory factor.
91. The isolated modified cell of claim 79, further comprising at least one third heterologous nucleic acid molecule that comprises a third polynucleotide sequence encoding a third polynucleotide.
92. The cell of claim 91, wherein the third polypeptide is a survival factor.
93. The cell of claim 92, wherein the survival factor is an anti-apoptotic factor.
94. The cell of claim 92, wherein the survival factor is Akt.
95. The cell of claim 92, wherein the survival factor is a reverse transcriptase.
96. The cell of claim 92, wherein the survival factor is telomerase.
97. The cell of claim 79, wherein the first heterologous nucleic acid sequence is under the regulatory control of a promoter.
98. The cell of claim 97, wherein the promoter is an inducible promoter.
99. The cell of claim 97, wherein the promoter is a constitutive promoter.
100. The cell of claim 97, wherein the promoter is a tissue-specific promoter.
101. The cell of claim 100, wherein the tissue-specific promoter enables expression of the first therapeutic molecule in a tissue chosen from liver, trachea, lung, heart, kidney, pancreas, brain, spinal cord, blood, bone marrow, joint, synovium, skin, GI tract, epithelial tissues, ovary, prostate, breast, spleen, bone, and cartilage.
102. The cell of claim 101, wherein the cell is differentiated into a tissue chosen from liver, trachea, lung, heart, kidney, pancreas, brain, spinal cord, blood, bone marrow, joint, synovium, skin, GI, epithelial tissues, ovary, prostate, breast, spleen, bone, and cartilage.
103. The cell of claim 97, wherein the promoter is heterologous to the first heterologous nucleic acid molecule.
104. The cell of claim 97, wherein the promoter is homologous to the first heterologous nucleic acid molecule.
105. A composition comprising the cell of claim 79 and a pharmaceutically acceptable carrier.

106. A kit comprising the composition of claim 106 and instructions for administration into an animal.

107. The kit of claim 106, wherein the animal is a human.

108. The cell of claim 87, wherein the second polynucleotide sequence is chosen from the appendices.

109. The cell of claim 108, wherein the first heterologous nucleic acid sequence encodes one member of a dimeric molecule and the second heterologous nucleic acid sequence encodes a second member of the dimeric molecule.

110. The cell of claim 109, wherein the dimeric molecule is a therapeutic molecule.

111. The cell of claim 109, wherein the dimeric molecule is an antibody.

112. The cell of claim 108, wherein the first heterologous nucleic acid sequence encodes an immunoglobulin heavy chain or a biologically active fragment thereof and the second heterologous nucleic acid sequence encodes an immunoglobulin light chain or a biologically active fragment thereof.

113. The cell of claim 111, wherein the antibody is an anti-inflammation antibody.

114. The cell of claim 111, wherein the antibody is an anti-cancer antibody.

115. A composition comprising the cell of claim 87 and a pharmaceutically acceptable carrier.

116. A kit comprising the composition of claim 115 and instructions for administration into an animal.

117. The kit of claim 116, wherein the animal is a human.

118. The isolated modified cell of claim 111, wherein the antibody is chosen from a monoclonal antibody, a single chain antibody, and a biologically active fragment thereof.

119. A method for treatment of a disease, disorder, syndrome, or condition in a subject comprising the steps of:

- (a) providing a composition comprising a plurality of isolated modified cells, wherein each modified cell comprises at least one first heterologous nucleic acid molecule, wherein the first heterologous nucleic acid molecule comprises a first polynucleotide sequence that encodes a first polypeptide, wherein the first polypeptide is chosen from the appendices; and
- (b) administering the composition to the subject.

120. The method of claim 119, wherein the disease, disorder, syndrome, or condition is a hyperproliferative disease, disorder, syndrome, or condition.

121. The method of claim 120, wherein the hyperproliferative disease, disorder, syndrome, or condition is chosen from cancer, psoriasis, and ulcerative colitis.

122. The method of claim 119, wherein the disease, disorder, syndrome, or condition is inflammation and the first polypeptide is an anti-inflammatory molecule.

123. The method of claim 119, wherein the disease, disorder, syndrome, or condition is ischemic heart disease and the first polypeptide is a vasodilator.

124. The method of claim 119, wherein the disease, disorder, syndrome, or condition is thrombosis and the first therapeutic molecule is an anti-coagulant.

125. The method of claim 119, wherein the disease, disorder, syndrome, or condition is immunological, and the first polypeptide is immunosuppressive.

126. The method of claim 119, wherein the disease, disorder, syndrome, or condition is immunological, and the first polypeptide enhances the immune response of the subject.

127. A method for treatment of a disease, disorder, syndrome, or condition in a subject comprising the steps of:

(a) providing a composition comprising a plurality of isolated modified cells of claim 87; and

(b) administering the composition to the subject.

128. The method of claim 127, wherein the disease, disorder, syndrome, or condition is a hyperproliferative disease, disorder, syndrome, or condition.

129. The method of claim 128, wherein the hyperproliferative disease, disorder, syndrome, or condition is chosen from among cancer, psoriasis, and ulcerative colitis.

130. The method of claim 129, wherein the hyperproliferative disease, disorder, syndrome, or condition is psoriasis.

131. A non-human animal deficient in a polypeptide chosen from SEQ. ID. NOS. 188 to 374.

132. The non-human animal of claim 131 in which a polynucleotide encoding the polypeptide is functionally defective or missing.

133. The non-human animal of claim 131, wherein the animal is a mouse.

134. The non-human animal of claim 131, wherein the non-human animal suffers from a disease, disorder, syndrome, or condition chosen from abnormal growth, hormonal imbalance, early aging, abnormal bone formation, immune disorder, cancer, diabetes, skin disease, and CNS disease.

135. Isolated tissues derived from the non-human animal of claim 131.
136. One or more cells derived from the non-human animal of claim 131.
137. The cell(s) of claim 136, wherein the cell(s) are chosen from stem cells, bone marrow cells, liver cells, pancreatic cells, skin cells, hematopoietic cells, kidney cells, heart cells, epithelial cells, neuronal cells, and lung cells.
138. The cell(s) of claim 136, wherein the cells are chosen from precursor cells, progenitor cells, and differentiated cells of the hematopoietic lineage.
139. The cell(s) of claim 136, wherein the cells are bone marrow cells, and the bone marrow cells are chosen from stromal cells and hematopoietic cells.
140. A non-human animal that over-expresses a polypeptide chosen from SEQ. ID. NOS. 188 to 374.
141. The non-human animal of claim 140, wherein the polypeptide is encoded by a nucleic acid molecule comprising an exogenous DNA sequence.
142. The non-human animal of claim 140, wherein the animal is a mouse.
143. The non-human animal of claim 140, wherein the non-human animal suffers from a disease, disorder, syndrome, or condition chosen from abnormal growth, hormonal imbalance, early aging, abnormal bone formation, immune disorder disease, cancer, diabetes, skin disease, and CNS disease.
144. One or more isolated tissues derived from the non-human animal of claim 140.
145. One or more cells derived from the non-human animal of claim 140.
146. The cell(s) of claim 145, wherein the cells are chosen from stem cells, bone marrow cells, liver cells, pancreatic cells, skin cells, hematopoietic cells, kidney cells, heart cells, epithelial cells, neuronal cells, and lung cells.
147. The cell(s) of claim 145, wherein the cells are chosen from precursor cells, progenitor cells, mesenchymal cells, stem cells, and differentiated cells of the hematopoietic lineage.
148. The cell(s) of claim 145, wherein the cells are bone marrow cells, and the bone marrow cells are chosen from stromal cells and hematopoietic cells.

DISCLOSURE OF THE INVENTION

[008] The present invention features an isolated polynucleotide that encodes a polypeptide. In some embodiments, the polypeptide has at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least

about 97%, at least about 98%, or at least about 99% amino acid sequence identity with an amino acid sequence derived from a polynucleotide sequence chosen from at least one nucleotide sequence according to SEQ. ID. NOS. 1-187 and 375-484. In some embodiments, the polypeptide has an amino acid sequence chosen from at least one amino acid sequence according to SEQ. ID. NOS. 188-374. In many embodiments, the polypeptide has at least one activity associated with the naturally occurring encoded polypeptide.

[009] In some embodiments, the polypeptide includes a signal peptide. In alternative embodiments, the polypeptide comprises a mature form of a protein, from which the signal peptide has been cleaved. In other embodiments, the polypeptide is a signal peptide. In a further aspect, the invention provides fragments of a polypeptide chosen from at least one amino acid sequence according to SEQ. ID. NOS. 188-374, where each fragment is an extracellular fragment of the polypeptide, or an extracellular fragment of the polypeptide minus the signal peptide. The invention provides an N-terminal fragment containing a Pfam domain and a C-terminal fragment containing a Pfam domain and either or both may be biologically active.

[010] In yet other embodiments, the polypeptides function as secreted proteins. In yet further embodiments, the polypeptides function as single-transmembrane proteins. In yet further embodiments, the polypeptides function as multiple-transmembrane proteins. In yet further embodiments, the polypeptides function as kinases, *e.g.*, protein kinases.

[011] The present invention features an isolated polynucleotide that hybridizes under stringent hybridization conditions to a coding region of at least one nucleotide sequence shown in SEQ. ID. NOS. 1-187, 375-484, or a complement thereof.

[012] The present invention features an isolated polynucleotide that shares at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97%, at least about 98%, at least about 99% nucleotide sequence identity with a nucleotide sequence of the coding region of at least one sequence shown in SEQ. ID. NOS. 1-187, 375-484, or a complement thereof. In some embodiments, a subject polynucleotide has the nucleotide sequence shown in at least one of SEQ. ID. NOS. 1-187, 375-484, or a coding region thereof.

[013] The present invention also features a vector, *e.g.*, a recombinant vector, that includes a subject polynucleotide, and a promoter that drives its expression. This vector can transform a host cell, and the present invention further features such host cells, *e.g.*, isolated *in*

vitro host cells, and *in vivo* host cells, that comprise a polynucleotide of the invention, or a recombinant vector of the invention.

[014] The present invention further features a library of polynucleotides, wherein at least one of the polynucleotides comprises the sequence information of a polynucleotide of the invention. In specific embodiments, the library is provided on a nucleic acid array. In some embodiments, the library is provided in computer-readable format.

[015] The present invention features a pair of isolated nucleic acid molecules, each from about 10 to about 200 nucleotides in length. The first nucleic acid molecule of the pair comprises a sequence of at least 10 contiguous nucleotides having 100% sequence identity to at least one nucleic acid sequence shown in SEQ. ID. NOS. 1-187 and 375-484. The second nucleic acid molecule of the pair comprises a sequence of at least 10 contiguous nucleotides having 100% sequence identity to the reverse complement of at least one nucleic acid sequence shown in SEQ. ID. NOS. 1-187 and 375-484. The sequence of said second nucleic acid molecule is located 3' of the nucleic acid sequence of the first nucleic acid molecule shown in SEQ. ID. NOS. 1-187 and 375-484. The pair of isolated nucleic acid molecules are useful in a polymerase chain reaction or in any other method known in the art to amplify a nucleic acid that has sequence identity to the sequences shown in SEQ. ID. NOS. 1-187 and 375-484, particularly when cDNA is used as a template.

[016] The invention features a method of determining the presence of a polynucleotide substantially identical to a polynucleotide sequence shown in the appendices, or a complement of such a nucleotide by providing its complement, allowing the polynucleotides to interact, and determining whether such interaction has occurred.

[017] The invention further features methods of regulating the expression of the subject polynucleotides and encoded polypeptides. The invention provides a method of inhibiting transcription or translation of a first polynucleotide encoding a first polypeptide of the invention by providing a second polynucleotide that hybridizes to the first polynucleotide, and allowing the first polynucleotide to contact and bind to the second polynucleotide. The second polynucleotide can be chosen from an antisense molecule, a ribozyme, and an interfering RNA (RNAi) molecule.

[018] The present invention further features an isolated polypeptide, *e.g.*, an isolated polypeptide encoded by a polynucleotide, and biologically active fragments of such polypeptide. In some embodiments, the polypeptide is a fusion protein. In some embodiments, the polypeptide has one or more amino acid substitutions, and/or insertions

and/or deletions, compared with at least one sequence shown in SEQ. ID. NOS. 188-374. In some embodiments, the polypeptide has an amino acid sequence derived from at least one nucleotide sequence shown in SEQ. ID. NOS. 1-187 and 375-484. In some embodiments, the polypeptide has an amino acid sequence substantially identical to at least one sequence shown in SEQ. ID. NOS. 188-374.

[019] The invention also provides a method of making a polypeptide of the invention by providing a nucleic acid molecule that comprises a polynucleotide sequence encoding a polypeptide of the invention, introducing the nucleic acid molecule into an expression system, and allowing the polypeptide to be produced.

[020] In some embodiments, the method involves *in vitro* cell-free transcription and/or translation. For example, the expression system can comprise a cell-free expression system, such as an *E. coli* system, a wheat germ extract system, a rabbit reticulocyte system, or a frog oocyte system.

[021] In certain other embodiments, the expression system can comprise a prokaryotic or eukaryotic cell, for example, a bacterial cell expression system, a fungal cell expression system, such as yeast or *Aspergillus*, a plant cell expression system, *e.g.*, a cereal plant, a tobacco plant, a tomato plant, or other edible plant, an insect cell expression system, such as SF9 of High Five cells, an amphibian cell expression system, a reptile cell expression system, a crustacean cell expression system, an avian cell expression system, a fish cell expression system, or a mammalian cell expression system, such as one using Chinese Hamster Ovary (CHO) cells. In some embodiments, the method involves culturing a subject host cell under conditions such that the subject polypeptide is produced by the host cells; and recovering the subject polypeptide from the culture, *e.g.*, from within the host cells, or from the culture medium. In further embodiments, the polypeptide can be produced *in vivo* in a multicellular animal or plant, comprising a polynucleotide encoding the subject polypeptide.

[022] The present invention further features a non-human animal injected with at least one polynucleotide comprising at least one nucleotide sequence chosen from SEQ. ID. NOS. 1-187 and 375-484, and/or at least one polypeptide comprising at least one amino acid sequence chosen from SEQ. ID. NOS. 188-374.

[023] The present invention further features an antibody that specifically recognizes, binds to, interferes with, or modulates the biological activity of a subject polypeptide or a fragment thereof. The present invention features an antibody that specifically inhibits

binding of a polypeptide to its ligand or substrate. It also features an antibody that specifically inhibits binding of a polypeptide as a substrate to another molecule.

[024] Another aspect of the present invention features a library of antibodies or fragments thereof, wherein at least one antibody or fragment thereof specifically binds to at least a portion of a polypeptide comprising an amino acid sequence according to SEQ. ID. NOS. 188-374 , and/or wherein at least one antibody or fragment thereof interferes with at least one activity of such polypeptide or fragment thereof. In certain embodiments, the antibody library comprises at least one antibody or fragment thereof that specifically inhibits binding of a subject polypeptide to its ligand or substrate, or that specifically inhibits binding of a subject polypeptide as a substrate to another molecule. The present invention also features corresponding polynucleotide libraries comprising at least one polynucleotide sequence that encodes an antibody or antibody fragment of the invention. In specific embodiments, the library is provided on a nucleic acid array or in computer-readable format.

[025] An antibody of the present invention may comprise a monoclonal antibody, polyclonal antibody, single chain antibody, intrabody, and active fragments of any of these. The active fragments include variable regions from either heavy chains or light chains. The antibody can comprise the backbone of a molecule with an immunoglobulin domain, *e.g.*, a fibronectin backbone, a T-cell receptor backbone, or a CTLA4 backbone.

[026] The present invention further features a targeting antibody, a neutralizing antibody, a stabilizing antibody, an enhancing antibody, an antibody agonist, an antibody antagonist, an antibody that promotes cellular endocytosis of a target antigen, a cytotoxic antibody, and an antibody that mediates antibody dependent cellular cytotoxicity (ADCC). The antibody that mediates ADCC can have a cytotoxic component, *e.g.*, a radioisotope, a radioactive molecule, a microbial toxin, a plant toxin, a chemotherapeutic agent, or a chemical substance, such as doxorubicin or cisplatin. The invention also features an inhibitory antibody, functioning to specifically inhibit the binding of a cognate polypeptide to its ligand or its substrate, or to specifically inhibit the binding of a cognate peptide as the substrate of another molecule.

[027] The antibodies of the present invention also encompass a human antibody, a non-human primate antibody, a monkey antibody, a non-primate animal antibody, *e.g.*, a rodent antibody, rat antibody, a mouse antibody, a hamster antibody, a guinea pig antibody, a chicken antibody, a cattle antibody, a sheep antibody, a goat antibody, a horse antibody,

porcine antibody, a cow antibody, a rabbit antibody, a cat antibody, or a dog antibody. It also features a humanized antibody, a primatized antibody, and a chimeric antibody.

[028] The antibodies of the invention can be produced *in vitro* or *in vivo*. For example, the present invention features an antibody produced in a cell-free expression system, a prokaryote expression system or a eukaryote expression system, as described herein.

[029] The invention further provides a host cell that can produce an antibody of the invention or a fragment thereof. The antibody may also be secreted by the cell. The host cell can be a hybridoma, or a prokaryotic or eukaryotic cell. The invention also provides a bacteriophage or other virus particle comprising an antibody of the invention, or a fragment thereof. The bacteriophage or other virus particle may display the antibody or fragment thereof on its surface, and the bacteriophage itself may exist within a bacterial cell. The antibody may also comprise a fusion protein with a viral or bacteriophage protein.

[030] The invention further provides transgenic multicellular organisms, *e.g.*, plants or non-human animals, as well as tissues or organs, comprising a polynucleotide sequence encoding a subject antibody or fragment thereof. The organism, tissues, or organs will generally comprise cells producing an antibody of the invention, or a fragment thereof.

[031] In another aspect, the present invention features a method of making an antibody by immunizing a host animal. In this method, a polypeptide or a fragment thereof, a polynucleotide encoding a polypeptide, or a polynucleotide encoding a fragment thereof, is introduced into an animal in a sufficient amount to elicit the generation of antibodies specific to the polypeptide or fragment thereof, and the resulting antibodies are recovered from the animal. The polypeptide can be encoded by a nucleic acid molecule comprising a nucleotide sequence chosen from at least one polynucleotide sequence according to SEQ. ID. NOS. 1-187 and 375-484. For example, the polypeptide may comprise at least one amino acid sequence chosen from SEQ. ID. NOS. 188-374.

[032] The invention thus also provides a non-human animal comprising an antibody of the invention. The animal can be a non-human primate, (*e.g.*, a monkey) a rodent (*e.g.*, a rat, a mouse, a hamster, a guinea pig), a chicken, cattle (*e.g.*, a cow), a sheep, a goat, a horse, a pig, a rabbit, a cat, or a dog.

[033] The present invention also features a method of making an antibody by isolating a spleen from an animal injected with a polypeptide or a fragment thereof, a polynucleotide encoding a polypeptide, or a polynucleotide encoding a fragment thereof, and

recovering antibodies from the spleen cells. Hybridomas can be made from the spleen cells, and hybridomas secreting specific antibodies can be selected.

[034] The present invention further features a method of making a polynucleotide library from spleen cells, and selecting a cDNA clone that produces specific antibodies, or fragments thereof. The cDNA clone or a fragment thereof can be expressed in an expression system that allows production of the antibody or a fragment thereof, as provided herein.

[035] The invention also provides a method for determining the presence or measuring the level of a polypeptide that specifically binds to an antibody of the invention. This method involves allowing the antibody to interact with a sample, and determining whether interaction between the antibody and any polypeptide in the sample has occurred. Antibodies that specifically bind to at least one subject polypeptide are useful in diagnostic assays, *e.g.*, to detect the presence of a subject polypeptide. Similarly, the invention features a method of determining the presence of an antibody to a polypeptide of the invention, by providing the polypeptide, allowing the antibody and the polypeptide to interact, and determining whether interaction has occurred.

[036] The present invention further features a method of identifying an agent that modulates the level of a subject polypeptide (or an mRNA encoding a subject polypeptide) in a cell. The method generally involves contacting a cell (*e.g.*, a eukaryotic cell) that produces the subject polypeptide with a test agent; and determining the effect, if any, of the test agent on the level of the polypeptide in the cell.

[037] The present invention further features a method of identifying an agent that modulates biological activity of a subject polypeptide. The methods generally involve contacting a subject polypeptide with a test agent; and determining the effect, if any, of the test agent on the activity of the polypeptide. In certain embodiments, the polypeptide is expressed on a cell surface. In certain embodiments, the agent or modulator is an antibody, for example, where an antibody binds to the polypeptide or affects its biological activity.

[038] The present invention further features biologically active agents (or modulators) identified using a method of the invention.

[039] The present invention also features a method of modulating biological activity using an agent selectable by the above methods. Briefly, the method of modulating biological activity comprises contacting the agent with a first human or a non-human host cell, thereby modulating the activity of the first host cell or a second host cell. In one example, contacting the agent with the first human or non-human host cell results in the

recruitment of a second host cell. The agent may be an antibody or antibody fragment of the invention.

[040] The modulation can comprise directly enhancing cell activity, indirectly enhancing cell activity, directly inhibiting cell activity, or indirectly inhibiting cell activity. The cell activity that is modulated can include transcription, translation, cell cycle control, signal transduction, intracellular trafficking, cell adhesion, cell mobility, proteolysis, ion transport, water transport, DNA repair, hydrolysis, lipase activity, polymerization using an RNA template or a DNA template, and nuclease activity. The modulation can result in cell death or apoptosis, or inhibition of cell death or apoptosis, as well as cell growth, cell proliferation, or cell survival, or inhibition of cell growth, cell proliferation, or cell survival; as well as mucosal preservation, inhibition of eicosanoid synthesis, or resistance to infection by viruses.

[041] Either the first or the second host cell can be a human or a non-human host cell. Either the first or the second host cell can be an immune cell, *e.g.*, a T cell, B cell, NK cell, dendritic cell, macrophage, muscle cell, stem cell, skin cell, fat cell, blood cell, brain cell, bone marrow cell, endothelial cell, retinal cell, bone cell, kidney cell, pancreatic cell, liver cell, spleen cell, prostate cell, cervical cell, ovarian cell, breast cell, tracheal cell, lung cell, liver cell, soft tissue cell, colorectal cell, other cell of the gastrointestinal tract, or a cancer cell.

[042] The invention also provides a method of diagnosing cancer, proliferative, inflammatory, immune, viral, bacterial, or metabolic disorder in a patient, by allowing an antibody specific for a polypeptide of the invention to contact a patient sample, and detecting specific binding between the antibody and any antigen in the sample to determine whether the subject has cancer, proliferative, inflammatory, immune, viral, bacterial, or metabolic disorder.

[043] The invention further provides a method of diagnosing cancer, proliferative, inflammatory, immune, viral, bacterial, or metabolic disorder in a patient, by allowing a polypeptide of the invention to contact a patient sample, and detecting specific binding between the polypeptide and any interacting molecule in the sample to determine whether the subject has cancer, proliferative, inflammatory, immune, viral, bacterial, or metabolic disorder.

[044] The invention also features a method of providing a polynucleotide, a polypeptide, or an agent of the invention, such as an antibody, to a subject by oral, buccal,

nasal, rectal, intraperitoneal, intradermal, transdermal, intratracheal, intrathecal, or parenteral administration, or otherwise by implantation or inhalation. For example, the polynucleotide, polypeptide or agent can be administered intranasally, intravenously, intra-arterially, intracardiacally, subcutaneously, intraperitoneally, transdermally, intraventricularly, or intracranially. The invention also provides a method for formulating a polynucleotide, polypeptide, or modulator composition, such as an antibody composition, for delivery by any of the routes of administration provided above, for example, for treatment of disorders. For example, the parenteral delivery can be via inhalation or implantation. The parenteral delivery can also be oral, intranasal, intraventricular, or intracranial.

[045] The present invention also features a pharmaceutical composition comprising a polynucleotide, polypeptide, or modulator of the invention and a carrier. The carrier can be a pharmaceutically acceptable carrier. The modulator can be obtainable by any methods of the invention, for example, the modulator can be an antibody or a fragment thereof. Further, oral formulations, preparations for injection, aerosol formulations, and suppositories can be prepared, each comprising the polynucleotide, polypeptide, or modulator composition. Further, nucleic acid compositions comprising polynucleotide sequences encoding the subject antibodies, or fragments thereof, can be prepared for administration to a subject.

[046] The invention also features a non-human animal injected with the polynucleotide, polypeptide, or modulator composition, for example the antibody composition. Again, the animal can be a non-human primate, (*e.g.*, a monkey) a rodent (*e.g.*, a rat, a mouse, a hamster, a guinea pig), a chicken, cattle (*e.g.*, a cow), a sheep, a goat, a horse, a pig, a rabbit, a cat, or a dog.

[047] In another aspect, the invention provides a method of treating a disorder in a subject needing or desiring such treatment, comprising administering a polynucleotide, polypeptide, or modulator of the invention to the subject. The subject can be a human or a non-human animal. The disorder can be cancer, proliferative, inflammatory, immune, metabolic, ulcerative, bacterial, or viral disorders.

[048] For example, the method of treatment may comprise administering an antibody composition with a first antibody that specifically binds to a first epitope of a first polypeptide or a fragment thereof, or that interferes with at least one activity of the first polypeptide or a fragment thereof, wherein the first polypeptide is encoded by a nucleic acid molecule comprising a nucleotide sequence chosen from SEQ. ID. NOS. 1-187 and 375-484, or any nucleic acid of the present invention. For example, the first polypeptide may comprise

an amino acid sequence chosen from SEQ. ID. NOS. 188-374. In certain embodiments, this method further comprises using a second antibody that binds specifically to or interferes with the activity of a second epitope of the first polypeptide or to a first epitope of a second polypeptide. The second polypeptide can be encoded by a nucleic acid molecule comprising a nucleotide sequence chosen from SEQ. ID. NOS. 1-187 and 375-484, or any nucleic acid of the present invention. For example, the second polypeptide may comprise an amino acid sequence chosen from SEQ. ID. NOS. 188-374. In certain embodiments, the antibody binds, or interferes with the activity of, at least one polypeptide fragment, wherein the fragment is an extracellular fragment of the polypeptide, or an extracellular fragment of the polypeptide minus the signal peptide, for the treatment, for example, of proliferative disorders, such as cancer.

[049] In other embodiments, the modulator may bind to a cell surface molecule that is over-expressed in the disorder. Further the modulator may be linked to an antibody of the invention. The antibody can be capable of initiating antibody dependent cell cytotoxicity, *e.g.*, where the antibody is in turn coupled to cytotoxic agents. This method is applicable when the disorder is cancer, another proliferative disorder, inflammatory, immune, bacterial, viral, or metabolic disorder, and the cell surface molecule is over-expressed in a cancer cell, diseased cell or virus-infected cell.

[050] The invention also provides a method for prophylactic or therapeutic treatment of a subject needing or desiring such treatment by providing a vaccine, that can be administered to the subject. The vaccine may comprise one or more of a polynucleotide, polypeptide, or modulator of the invention, for example an antibody vaccine composition, a polypeptide vaccine composition, or a polynucleotide vaccine composition, useful for treating cancer, proliferative, inflammatory, immune, metabolic, bacterial, or viral disorders.

[051] For example, the vaccine can be a cancer vaccine, and the polypeptide can concomitantly be a cancer antigen. The vaccine may be an anti-inflammatory vaccine, and the polypeptide can concomitantly be an inflammation-related antigen. The vaccine may be a viral vaccine, and the polypeptide can concomitantly be a viral antigen. In some embodiments, the vaccine comprises a polypeptide fragment, comprising at least one extracellular fragment of a polypeptide of the invention, and/or at least one extracellular fragment of a polypeptide of the invention minus the signal peptide, for the treatment, for example, of proliferative disorders, such as cancer. In certain embodiments, the vaccine comprises a polynucleotide encoding one or more such fragments, administered for the

treatment, for example, of proliferative disorders, such as cancer. Further, the vaccine can be administered with or without an adjuvant.

[052] In another aspect, the invention provides a method for gene therapy by providing a polynucleotide comprising a nucleic acid molecule encoding a polypeptide, such as an antibody of the invention, and administering the polynucleotide to a subject needing or desiring such treatment.

[053] The invention further provides a kit comprising one or more of a polynucleotide, polypeptide, or modulator composition, such as an antibody composition, which may include instructions for its use. Such kits are useful in diagnostic applications, for example, to detect the presence and/or level of a polypeptide in a biological sample by specific antibody interaction.

[054] Additional objects and advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or may be learned by practice of the invention. The objects and advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. Moreover, advantages described in the body of the specification, if not included in the claims, are not per se limitations to the claimed invention.

Brief Description of the Tables and Appendices

[055] Each sequence shown in Tables 1-3 is identified by an internal reference number (FP ID). Table 1 lists the sequences of the invention. Each is identified by its FP ID number, a SEQ. ID. NO. corresponding to the nucleotide coding sequence (SEQ. ID. NO. (N1), a SEQ. ID. NO. corresponding to the encoded polypeptide sequence (SEQ. ID. NO. (P1), and a SEQ. ID. NO. corresponding to the entire nucleotide sequence (SEQ. ID. NO. (N0). Each is also identified by a Clone ID designation that lists each novel clone of the invention.

[056] Table 2 describes structural characteristics of the polypeptides of the invention encoded by these novel clones. In addition to listing the FP ID and Clone ID, it specifies the predicted number of amino acid residues in the polypeptide (Predicted Protein Length). Table 2 also specifies the result of an algorithm that predicts whether a sequence is secreted (Tree Vote). This algorithm is constructed on the basis of a number of attributes that include hydrophobicity, two-dimensional structure, prediction of signal sequence cleavage site, and other parameters, and predicts whether the listed sequences are secreted polypeptides or nucleotides related to secreted polypeptides. Table 2 also specifies the positions of the amino

acid residues that comprise signal peptides (Signal Peptide Coords.), as well as the coordinates of the amino acid residues comprising the protein after the signal peptide sequence has been cleaved to produce a mature protein (Mature Protein Coords.). Additionally, Table 2 specifies the coordinates of an alternate form of a mature protein. In instances where the mature protein start residue overlaps the signal peptide end residue, some of the amino acid residues may be cleaved off such that the mature protein does not start at the next amino acid residue from the signal peptides, resulting in the alternative mature protein coordinate (Alternative Mature Protein Coords.). Table 2 specifies the number of Transmembrane domains of each of the polypeptides of the invention (TM). It specifies the positions of the amino acid residues that comprise the transmembrane domains (TM Coords.) and the positions of the amino acids that do not pass through the membrane (Non-TM Coords.). Finally, Table 2 specifies protein family (Pfam) classifications for some of these novel human cDNA clones.

[057] Table 3 describes the characteristics of the protein in the public National Center for Information Biotechnology (NCBI) database displaying the greatest degree of similarity to polypeptides encoded by each novel human cDNA clone of the invention. The NCBI protein with the greatest homology to each Clone ID is described by its NCBI accession number (Top Hit Accession ID), and by the NCBI's annotation of that sequence (Top Hit Annotation). The percent identity of the Five Prime protein with its corresponding NCBI protein is listed (Top Hit % ID). Table 3 also describes the characteristics of the human protein in the NCBI database with the greatest degree of similarity to polypeptides encoded by each novel human cDNA clone of the invention. The corresponding NCBI protein is described by its NCBI accession number (Top Human Hit Accession ID) and by the NCBI's annotation of that sequence (Top Human Hit Annotation). Finally, the percent identity of the Five Prime protein with the NCBI protein is listed (Top Human Hit % ID).

[058] Appendix A provides nucleotide sequences of the encoding regions of the subject nucleic acids. Appendix B provides amino acid sequences of the subject polypeptides, which correspond to the encoding sequences in Appendix A. Appendix C provides nucleotide sequences of the DNA sequences used to find the encoding sequences of the subject nucleic acids. The appendices are provided as a paper copy.

Definitions

[059] "Related sequences" include nucleotide and amino acid sequences that are involved in the function of their referent. For example, "receptor-related sequences" include

all sequences that are involved in receptor function. This includes, but is not limited to, sequences that are involved in receptor synthesis, receptor regulation, receptor effector function, and receptor degradation. "Related sequences" also encompass complementary nucleic acid sequences, and biologically active fragments of nucleic acid and amino acid sequences.

[060] The terms "polynucleotide," "nucleotide," "nucleic acid," "polynucleic molecule," "nucleotide molecule," "nucleic acid molecule," "nucleic acid sequence," "polynucleotide sequence," and "nucleotide sequence" are used interchangeably herein to refer to polymeric forms of nucleotides of any length. The polynucleotides can contain deoxyribonucleotides, ribonucleotides, and/or their analogs or derivatives. For example, nucleic acids can be naturally occurring DNA or RNA, or can be synthetic analogs, as known in the art. The terms also encompass genomic DNA, genes, gene fragments, exons, introns, regulatory sequences or regulatory elements (such as promoters, enhancers, initiation and termination regions, other control regions, expression regulatory factors, and expression controls), DNA comprising one or more single-nucleotide polymorphisms (SNPs), allelic variants, isolated DNA of any sequence, and cDNA. The terms also encompass mRNA, tRNA, rRNA, ribozymes, splice variants, antisense RNA, antisense conjugates, RNAi, and isolated RNA of any sequence. The terms also encompass recombinant polynucleotides, heterologous polynucleotides, branched polynucleotides, labeled polynucleotides, hybrid DNA/RNA, polynucleotide constructs, vectors comprising the subject nucleic acids, nucleic acid probes, primers, and primer pairs. The polynucleotides can comprise modified nucleic acid molecules, with alterations in the backbone, sugars, or heterocyclic bases, such as methylated nucleic acid molecules, peptide nucleic acids, and nucleic acid molecule analogs, which may be suitable as, for example, probes if they demonstrate superior stability and/or binding affinity under assay conditions. Analogs of purines and pyrimidines, including radiolabeled and fluorescent analogs, are known in the art. The polynucleotides can have any three-dimensional structure, and can perform any function, known or as yet unknown. The terms also encompass single-stranded, double-stranded and triple helical molecules that are either DNA, RNA, or hybrid DNA/RNA and that may encode a full-length gene or a biologically active fragment thereof. Biologically active fragments of polynucleotides can encode the polypeptides herein, as well as anti-sense and RNAi molecules. Thus, the full length polynucleotides herein may be treated with enzymes, such as Dicer, to generate a library of short RNAi fragments which are within the scope of the present invention.

[061] The novel polynucleotides herein include those shown in the Tables, SEQ. ID. NOS. 1-187 and 375-484, as well as those that encode the polypeptides of SEQ. ID. NOS. 188-374, and biologically active fragments thereof. The polynucleotides also include modified, labeled, and degenerate variants of the nucleic acid sequences, as well as nucleic acid sequences that are substantially similar or homologous to nucleic acids encoding the subject proteins.

[062] A "biologically active" entity, or an entity having "biological activity," is one having structural, regulatory, or biochemical functions of a naturally occurring molecule or any function related to or associated with a metabolic or physiological process. Biologically active polynucleotide fragments are those exhibiting activity similar, but not necessarily identical, to an activity of a polynucleotide of the present invention. The biological activity can include an improved desired activity, or a decreased undesirable activity. For example, an entity demonstrates biological activity when it participates in a molecular interaction with another molecule, or when it has therapeutic value in alleviating a disease condition, or when it has prophylactic value in inducing an immune response to the molecule, or when it has diagnostic value in determining the presence of the molecule, such as a biologically active fragment of a polynucleotide that can be detected as unique for the polynucleotide molecule, or that can be used as a primer in PCR.

[063] The term "degenerate variant" of a nucleic acid sequence refers to all nucleic acid sequences that can be directly translated, according to the standard genetic code, to provide an amino acid sequence identical to that translated from a reference nucleic acid sequence.

[064] The term "gene" or "genomic sequence" as used herein is an open reading frame encoding specific proteins and polypeptides, for example, an mRNA, cDNA, or genomic DNA, and also may or may not include intervening introns, or adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression up to about 20 kb beyond the coding region, and possibly further in either direction. A gene can be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome.

[065] The term "transgene" as used herein is a nucleic acid sequence that is incorporated into a transgenic organism. A "transgene" can contain one or more transcriptional regulatory sequences, and other sequences, such as introns, that may be useful for expressing or secreting the nucleic acid or fusion protein it encodes.

[066] The term "cDNA" as used herein is intended to include all nucleic acids that share the sequence elements of mature mRNA species, where sequence elements are exons and 3' and 5' non-coding regions. Generally, mRNA species have contiguous exons, the intervening introns having been removed by nuclear RNA splicing to create a continuous open reading frame encoding a protein.

[067] The term "splice variant" refers to all types of RNAs transcribed from a given gene that when processed collectively encode plural protein isoforms. The term "alternative splicing" and related terms refer to all types of RNA processing that lead to expression of plural protein isoforms from a single gene. Some genes are first transcribed as long mRNA precursors that are then shortened by a series of processing steps to produce the mature mRNA molecule. One of these steps is RNA splicing, in which the intron sequences are removed from the mRNA precursor. A cell can splice the primary transcript in different ways, making different "splice variants," and thereby making different polypeptide chains from the same gene, or from the same mRNA molecule. Splice variants can include, for example, exon insertions, exon extensions, exon truncations, exon deletions, alternatives in the 5' untranslated region and alternatives in the 3' untranslated region.

[068] "Oligonucleotide" may generally refer to polynucleotides of between about 5 and about 100 nucleotides of single-or double-stranded nucleic acids. For the purposes of this disclosure, there is no upper limit to the length of an oligonucleotide. Oligonucleotides are also known as oligomers or oligos and can be isolated from genes, or chemically synthesized by methods known in the art.

[069] "Nucleic acid composition" as used herein is a composition comprising a nucleic acid sequence, including one having an open reading frame that encodes a polypeptide and is capable, under appropriate conditions, of being expressed as a polypeptide. The term includes, for example, vectors, including plasmids, cosmids, viral vectors (*e.g.*, retrovirus vectors such as lentivirus, adenovirus, and the like), human, yeast, bacterial, P1-derived artificial chromosomes (HAC's, YAC's, BAC's, PAC's, etc), and mini-chromosomes, *in vitro* host cells, *in vivo* host cells, tissues, organs, allogenic or congenic grafts or transplants, multicellular organisms, and chimeric, genetically modified, or transgenic animals comprising a subject nucleic acid sequence.

[070] An "isolated," "purified," or "substantially isolated" polynucleotide, or a polynucleotide in "substantially pure form," in "substantially purified form," in "substantial purity," or as an "isolate," is one that is substantially free of the sequences with which it is

associated in nature, or other nucleic acid sequences that do not include a sequence or fragment of the subject polynucleotides. By substantially free is meant that less than about 90%, less than about 80%, less than about 70%, less than about 60%, or less than about 50% of the composition is made up of materials other than the isolated polynucleotide. For example, the isolated polynucleotide is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 97%, or at least about 99% free of the materials with which it is associated in nature. For example, an isolated polynucleotide may be present in a composition wherein at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 97%, at least about 99% of the total macromolecules (for example, polypeptides, fragments thereof, polynucleotides, fragments thereof, lipids, polysaccharides, and oligosaccharides) in the composition is the isolated polynucleotide. Where at least about 99% of the total macromolecules is the isolated polynucleotide, the polynucleotide is at least about 99% pure, and the composition comprises less than about 1% contaminant. As used herein, an "isolated," "purified" or "substantially isolated" polynucleotide, or a polynucleotide in "substantially pure form," in "substantially purified form," in "substantial purity," or as an "isolate," also refers to recombinant polynucleotides, modified, degenerate and homologous polynucleotides, and chemically synthesized polynucleotides, which, by virtue of origin or manipulation, are not associated with all or a portion of a polynucleotide with which it is associated in nature, are linked to a polynucleotide other than that to which it is linked in nature, or do not occur in nature. For example, the subject polynucleotides are generally provided as other than on an intact chromosome, and recombinant embodiments are typically flanked by one or more nucleotides not normally associated with the subject polynucleotide on a naturally-occurring chromosome.

[071] The terms "polypeptide," "peptide," and "protein," used interchangeably herein, refer to a polymeric form of amino acids of any length, which can include naturally-occurring amino acids, coded and non-coded amino acids, chemically or biochemically modified, derivatized, or designer amino acids, amino acid analogs, peptidomimetics, and depsipeptides, and polypeptides having modified, cyclic, bicyclic, depsicyclic, or depsibicyclic peptide backbones. The term includes single chain protein as well as multimers. The term also includes conjugated proteins, fusion proteins, including, but not limited to, GST fusion proteins, fusion proteins with a heterologous amino acid sequence, fusion proteins with heterologous and homologous leader sequences, fusion proteins with or

without N-terminal methionine residues, pegylated proteins, and immunologically tagged proteins. Also included in this term are variations of naturally occurring proteins, where such variations are homologous or substantially similar to the naturally occurring protein, as well as corresponding homologs from different species. Variants of polypeptide sequences include insertions, additions, deletions, or substitutions compared with the subject polypeptides. The term also includes peptide aptamers.

[072] The novel polypeptides herein include amino acid sequences encoded by an open reading frame (ORF) as shown in SEQ. ID. NOS. 188-374, described in greater detail below, including the full length protein and fragments thereof, particularly biologically active fragments and/or fragments corresponding to functional domains, *e.g.*, a signal peptide or leader sequence, an enzyme active site, including a cleavage site and an enzyme catalytic site, a domain for interaction with other protein(s), a domain for binding DNA, a regulatory domain, a consensus domain that is shared with other members of the same protein family, such as a kinase family or an immunoglobulin family; an extracellular domain that may act as a target for antibody production or that may be cleaved to become a soluble receptor or a ligand for a receptor; an intracellular fragment of a transmembrane protein that participates in signal transduction; a transmembrane domain of a transmembrane protein that may facilitate water or ion transport; a sequence associated with cell survival and/or cell proliferation; a sequence associated with cell cycle arrest, DNA repair and/or apoptosis; a sequence associated with a disease or disease prognosis, including types of cancer, degenerative disease, inflammatory disease, immunological disease, genetic disease, metabolic disease, and/or bacterial or viral infection; and including fusions of the subject polypeptides to other proteins or parts thereof; modifications of the subject polypeptide, *e.g.*, comprising modified, derivatized, or designer amino acids, modified peptide backbones, and/or immunological tags; as well as intra- and inter-species homologs of the subject polypeptides.

[073] As noted above, a "biologically active" entity, or an entity having "biological activity," is one having structural, regulatory, or biochemical functions of a naturally occurring molecule or any function related to or associated with a metabolic or physiological process. Biologically active polypeptide fragments are those exhibiting activity similar, but not necessarily identical, to an activity of a polypeptide of the present invention. The biological activity can include an improved desired activity, or a decreased undesirable activity. For example, an entity demonstrates biological activity when it participates in a molecular interaction with another molecule, or when it has therapeutic value in alleviating a

disease condition, or when it has prophylactic value in inducing an immune response to the molecule, or when it has diagnostic value in determining the presence of the molecule. A biologically active polypeptide or fragment thereof includes one that can participate in a biological reaction, for example, as a transcription factor that combines with other transcription factors for initiation of transcription, or that can serve as an epitope or immunogen to stimulate an immune response, such as production of antibodies, or that can transport molecules into or out of cells, or that can perform a catalytic activity, for example polymerization or nuclease activity, or that can participate in signal transduction by binding to receptors, proteins, or nucleic acids, activating enzymes or substrates.

[074] A "signal peptide," or a "leader sequence," comprises a sequence of amino acid residues, typically, at the N terminus of a polypeptide, which directs the intracellular trafficking of the polypeptide. Polypeptides that contain a signal peptide or leader sequence typically also contain a signal peptide or leader sequence cleavage site. Such polypeptides, after cleavage at the cleavage sites, generate mature polypeptides, for example, after extracellular secretion or after being directed to the appropriate intracellular compartment.

[075] "Depsipeptides" are compounds containing a sequence of at least two alpha-amino acids and at least one alpha-hydroxy carboxylic acid, which are bound through at least one normal peptide link and ester links, derived from the hydroxy carboxylic acids. "Linear depsipeptides" can comprise rings formed through S-S bridges, or through an hydroxy or a mercapto group of an hydroxy-, or mercapto-amino acid and the carboxyl group of another amino- or hydroxy-acid but do not comprise rings formed only through peptide or ester links derived from hydroxy carboxylic acids. "Cyclic depsipeptides" are peptides containing at least one ring formed only through peptide or ester links, derived from hydroxy carboxylic acids.

[076] An "isolated," "purified," or "substantially isolated" polypeptide, or a polypeptide in "substantially pure form," in "substantially purified form," in "substantial purity," or as an "isolate," is one that is substantially free of the materials with which it is associated in nature or other polypeptide sequences that do not include a sequence or fragment of the subject polypeptides. By substantially free is meant that less than about 90%, less than about 80%, less than about 70%, less than about 60%, or less than about 50% of the composition is made up of materials other than the isolated polypeptide. For example, the isolated polypeptide is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 97%, or at least about 99%

free of the materials with which it is associated in nature. For example, an isolated polypeptide may be present in a composition wherein at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 97%, or at least about 99% of the total macromolecules (for example, polypeptides, fragments thereof, polynucleotides, fragments thereof, lipids, polysaccharides, and oligosaccharides) in the composition is the isolated polypeptide. Where at least about 99% of the total macromolecules is the isolated polypeptide, the polypeptide is at least about 99% pure, and the composition comprises less than about 1% contaminant. As used herein, an "isolated," "purified," or "substantially isolated" polypeptide, or a polypeptide in "substantially pure form," in "substantially purified form," in "substantial purity," or as an "isolate," also refers to recombinant polypeptides, modified, tagged and fusion polypeptides, and chemically synthesized polypeptides, which by virtue of origin or manipulation, are not associated with all or a portion of the materials with which they are associated in nature, are linked to molecules other than that to which they are linked in nature, or do not occur in nature.

[077] Detection methods of the invention can be qualitative or quantitative. Thus, as used herein, the terms "detection," "identification," "determination," and the like, refer to both qualitative and quantitative determinations, and include "measuring." For example, detection methods include methods for detecting the presence and/or level of polynucleotide or polypeptide in a biological sample, and methods for detecting the presence and/or level of biological activity of polynucleotide or polypeptide in a sample.

[078] As used herein, the term "array" or "microarray" may be used interchangeably and refers to a collection of plural biological molecules such as nucleic acids, polypeptides, or antibodies, having locatable addresses that may be separately detectable. Generally, "microarray" encompasses use of sub microgram quantities of biological molecules. The biological molecules may be affixed to a substrate or may be in solution or suspension. The substrate can be porous or solid, planar or non-planar, unitary or distributed, such as a glass slide, a 96 well plate, with or without the use of microbeads or nanobeads. As such, the term "microarray" includes all of the devices referred to as microarrays in Schena, 1999; Bassett et al., 1999; Bowtell, 1999; Brown and Botstein, 1999; Chakravarti, 1999; Cheung et al., 1999; Cole et al., 1999; Collins, 1999; Debouck and Goodfellow, 1999; Duggan et al., 1999; Hacia, 1999; Lander, 1999; Lipshutz et al., 1999; Southern, et al., 1999; Schena, 2000; Brenner et al, 2000; Lander, 2001; Steinhaur et al., 2002; and Espejo et al, 2002. Nucleic acid microarrays

include both oligonucleotide arrays (DNA chips) containing expressed sequence tags ("ESTs") and arrays of larger DNA sequences representing a plurality of genes bound to the substrate, either one of which can be used for hybridization studies. Protein and antibody microarrays include arrays of polypeptides or proteins, including but not limited to, polypeptides or proteins obtained by purification, fusion proteins, and antibodies, and can be used for specific binding studies (Zhu and Snyder, 2003; Houseman et al., 2002; Schaeferling et al., 2002; Weng et al., 2002; Winssinger et al., 2002; Zhu et al., 2001; Zhu et al. 2001; and MacBeath and Schreiber, 2000).

[079] A "nucleic acid hybridization reaction" is one in which single strands of DNA or RNA randomly collide with one another, and bind to each other only when their nucleotide sequences have some degree of complementarity. The solvent and temperature conditions can be varied in the reactions to modulate the extent to which the molecules can bind to one another. Hybridization reactions can be performed under different conditions of "stringency." The "stringency" of a hybridization reaction as used herein refers to the conditions (*e.g.*, solvent and temperature conditions) under which two nucleic acid strands will either pair or fail to pair to form a "hybrid" helix.

[080] " T_m " is the temperature in degrees Celsius at which 50% of a polynucleotide duplex made of complementary strands of nucleic acids that are hydrogen bonded in an anti-parallel direction by Watson-Crick base pairing dissociate into single strands under conditions of the hybridization reaction. T_m can be predicted according to a standard formula, such as: $T_m = 81.5 + 16.6 \log[X^+] + 0.41 (\%G/C) - 0.61 (\%F) - 600/L$, where $[X^+]$ is the cation concentration (usually sodium ion, Na^+) in mol/L; ($\%G/C$) is the number of G and C residues as a percentage of total residues in the duplex; ($\%F$) is the percent formamide in solution (wt/vol); and L is the number of nucleotides in each strand of the paired nucleic acids.

[081] A "buffer" is a system that tends to resist change in pH when a given increment of hydrogen ion or hydroxide ion is added. Buffered solutions contain conjugate acid-base pairs. Any conventional buffer can be used with the inventions herein including but not limited to, for example, Tris, phosphate, imidazole, and bicarbonate.

[082] A "library" of polynucleotides comprises a collection of sequence information of a plurality of polynucleotide sequences, which information is provided in either biochemical form (*e.g.*, as a collection of polynucleotide molecules), or in electronic form

(*e.g.*, as a collection of polynucleotide sequences stored in a computer-readable form, as in a computer-based system, a computer data file, and/or as part of a computer program).

[083] A "library" of polypeptides comprises a collection of sequence information of a plurality of polypeptide sequences, which information is provided in, *e.g.*, a collection of polypeptide sequences stored in a computer-readable form, as in a computer-based system, a computer data file, and/or as part of a computer program.

[084] "Media" refers to a manufacture, other than an isolated nucleic acid molecule, that contains the sequence information of the present invention. Such a manufacture provides the genome sequence or a subset thereof in a form that can be examined by means not directly applicable to the sequence as it exists in a nucleic acid, *e.g.*, with computer-readable media comprising data storage structures. Such media include, but are not limited to: magnetic storage media, such as a floppy disc, a hard disc storage medium, and a magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media.

[085] "Recorded" refers to a process for storing information on computer readable media, using any such methods as known in the art.

[086] As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. The data storage means can comprise any manufacture comprising a recording of the present sequence information as described above, or a memory access means that can access such a manufacture.

[087] "Search means" refers to one or more programs implemented on the computer-based system, to compare a target sequence or target structural motif, or expression levels of a polynucleotide in a sample, with the stored sequence information. A variety of known algorithms are publicly known and commercially available, *e.g.*, MacPattern (EMBL), BLAST, BLASTN and BLASTX (NCBI), gapped BLAST, BLAZE, the Wise package, FASTX, Clustalw, FASTA, FASTA3, Align0, Toffee, BestFit, FastDB, and TeraBLAST (TimeLogic, Crystal Bay, Nevada). Search means can be used to identify fragments or regions of the genome that match a particular target sequence or target motif, for example,

based on sequence similarity, for example, to identify open reading frames (ORFs) within the genome that contain homology to ORFs from other organisms.

[088] "Sequence similarity," "sequence homology," "homology," "sequence identity," and "percent sequence identity," used interchangeably herein, describe the degree of relatedness between two polynucleotide or polypeptide sequences. In general, "identity" means the exact match-up of two or more nucleotide sequences or two or more amino acid sequences, where the nucleotide or amino acids being compared are the same. Also, in general, "similarity" or "homology" means the exact match-up of two or more nucleotide sequences or two or more amino acid sequences, where the nucleotide or amino acids being compared are either the same or possess similar chemical and/or physical properties. The terms also refer to the percentage of the "aligned" bases (for the polynucleotides) or amino acid residues (for the polypeptides) that are identical when the sequences are aligned. Sequences can be aligned in a number of different ways and sequence similarity can be determined in a number of different ways. For example, the bases or amino acid residues of one sequence can be aligned to a gap in the other sequence, or they can be aligned only to another base or amino acid residue in the other sequence. A gap can range anywhere from one nucleotide, base, or amino acid residue to multiple exons in length, up to any number of nucleotides or amino acid residues. Further, sequences can be aligned such that nucleotides (or bases) align with nucleotides, nucleotides align with amino acid residues, or amino acid residues align with amino acid residues.

[089] By "homolog" is meant a protein having at least about 35%, at least about 40%, at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, or at least about 95%, or higher, amino acid sequence identity to the reference polypeptide, as measured with the "GAP" program (part of the Wisconsin Sequence Analysis Package available through the Genetics Computer Group, Inc. (Madison WI)), where the parameters are: Gap weight:12; length weight:4. In many embodiments of interest, homology will be at least about 75%, at least about 80%, or at least 85%, where in certain embodiments of interest, homology will be as high as about 90%.

[090] A "target sequence" can be any polynucleotide or amino acid sequence of six or more contiguous nucleotides or two or more amino acids, for example, from about 5 or from about 10 to about 100 amino acids, or from about 15 or from about 30 to about 300 nucleotides. A variety of comparing means can be used to accomplish comparison of sequence information from a sample (*e.g.*, to analyze target sequences, target motifs, or

relative expression levels) with the data storage means. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer based systems of the present invention to accomplish comparison of target sequences and motifs. Computer programs to analyze expression levels in a sample and in controls are also known in the art. A "target sequence" includes an "antibody target sequence," which refers to an amino acid sequence that can be used as an immunogen for injection into animals for production of antibodies or for screening against a phage display or antibody library for identification of binding partners.

[091] A "target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration that is formed upon the folding of the target motif, or on consensus sequences of regulatory or active sites. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, hairpin structures, promoter sequences, and other expression elements such as binding sites for transcription factors.

[092] A "matrix" is a geometric network of antibody molecules and their antigens, as found in immunoprecipitation and flocculation reactions. An antibody matrix can exist in solution or on a solid phase support.

[093] The term "binds specifically," in the context of antibody binding, refers to high avidity and/or high affinity binding of an antibody to a specific polypeptide, or more accurately, to an epitope of a specific polypeptide. Antibody binding to such epitope on a polypeptide can be stronger than binding of the same antibody to any other epitopes, particularly other epitopes that can be present in molecules in association with, or in the same sample as the polypeptide of interest. For example, when an antibody binds more strongly to one epitope than to another, adjusting the binding conditions can result in antibody binding almost exclusively to the specific epitope and not to any other epitopes on the same polypeptide, and not to any other polypeptide, which does not comprise the epitope. Antibodies that bind specifically to a subject polypeptide may be capable of binding other polypeptides at a weak, yet detectable, level (*e.g.*, 10% or less of the binding shown to the polypeptide of interest). Such weak binding, or background binding, is readily discernible from the specific antibody binding to a subject polypeptide, *e.g.*, by use of appropriate

controls. In general, antibodies of the invention bind to a specific polypeptide with a binding affinity of 10^{-7} M or greater (*e.g.*, 10^{-8} M, 10^{-9} M, 10^{-10} , 10^{-11} , etc.).

[094] A recombinant vector or construct that includes a nucleic acid of the invention is useful for propagating a nucleic acid in a host cell; such vectors are known as "cloning vectors." Vectors can transfer nucleic acid between host cells derived from disparate organisms; these are known in the art as "shuttle vectors." Vectors can also insert a subject nucleic acid into a host cell's chromosome; these are known in the art as "insertion vectors." Vectors can express either sense or antisense RNA transcripts of the invention *in vitro* (*e.g.*, in a cell-free system or within an *in vitro* cultured host cell) or *in vivo* (*e.g.*, in a multicellular plant or animal); these are known in the art as "expression vectors," which can be part of an expression system. Expression vectors can also produce a subject antibody.

[095] The term "host cell" includes an individual cell, cell line, cell culture, or *in vivo* cell, which can be or has been a recipient of any polynucleotides or polypeptides of the invention, for example, a recombinant vector, an isolated polynucleotide, antibody or fusion protein. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology, physiology, or in total DNA, RNA, or polypeptide complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or change. Host cells can be prokaryotic or eukaryotic, including mammalian, insect, amphibian, reptile, crustacean, avian, fish, plant and fungal cells. A host cell includes cells transformed, transfected, transduced, or infected *in vivo* or *in vitro* with a polynucleotide of the invention, for example, a recombinant vector. A host cell which comprises a recombinant vector of the invention may be called a "recombinant host cell."

[096] "Biological sample," "patient sample," "clinical sample" "sample," or "biological specimen," used interchangeably herein, encompasses a variety of sample types obtained from an individual, including biological fluids such as blood, serum, plasma, urine, cerebrospinal fluid, tears, saliva, lymph, dialysis fluid, lavage fluid, semen, and other liquid samples or tissues of biological origin. It includes tissue samples and tissue cultures or cells derived therefrom and the progeny thereof, including cells in culture, cell supernatants, and cell lysates. It includes organ or tissue culture derived fluids, tissue biopsy samples, tumor biopsy samples, stool samples, and fluids extracted from physiological tissues. Cells dissociated from solid tissues, tissue sections, and cell lysates are included. The definition also includes samples that have been manipulated in any way after their procurement, such as by treatment with reagents, solubilization, or enrichment for certain components, such as

polynucleotides or polypeptides. Also included in the term are derivatives and fractions of biological samples. A biological sample can be used in a diagnostic, monitoring, or screening assay.

[097] The terms "individual," "host," "patient," and "subject," used interchangeably herein, refer to a mammal, including, but not limited to, murines, simians, humans, felines, canines, equines, bovines, porcines, ovines, caprines, mammalian farm animals, mammalian sport animals, and mammalian pets. "Mammals" or "mammalian," are used broadly to describe organisms which are within the class mammalia, including the orders carnivore (*e.g.*, dogs and cats), rodentia (*e.g.*, mice, guinea pigs, and rats), and other mammals, including cattle, goats, sheep, cows, horses, rabbits, and pigs, and primates (*e.g.*, humans, chimpanzees, and monkeys).

[098] The terms "agent," "substance," "modulator," and "compound" are used interchangeably herein. These terms refer to a substance that binds to or modulates a level or activity of a subject polypeptide or a level of mRNA encoding a subject protein or nucleic acid, or that modulates the activity of a cell containing the subject protein or nucleic acid. Where the agent modulates a level of mRNA encoding a subject protein, agents include ribozymes, antisense, and RNAi molecules. Where the agent is a substance that modulates a level of activity of a subject polypeptide, agents include antibodies specific for the subject polypeptide, peptide aptamers, small molecules, agents that bind a ligand-binding site in a subject polypeptide, and the like. Antibody agents include antibodies that specifically bind a subject polypeptide and activate the polypeptide, such as receptor-ligand binding that initiates signal transduction; antibodies that specifically bind a subject polypeptide and inhibit binding of another molecule to the polypeptide, thus preventing activation of a signal transduction pathway; antibodies that bind a subject polypeptide to modulate transcription; antibodies that bind a subject polypeptide to modulate translation; as well as antibodies that bind a subject polypeptide on the surface of a cell to initiate ADCC or to initiate cell killing or cell growth. Small molecule agents include those that bind the polypeptide to modulate activity of the polypeptide or cell containing the polypeptide in a similar fashion. The term "agent" also refers to substances that modulate a condition or disorder associated with a subject polynucleotide or polypeptide. Such agents include subject polynucleotides themselves, subject polypeptides themselves, and the like. Agents may be chosen from amongst candidate agents, as defined below.

[009] The terms "candidate agent," "subject agent," or "test agent," used interchangeably herein, encompass numerous chemical classes, typically synthetic, semi-synthetic, or naturally occurring inorganic or organic molecules, small molecules, or macromolecular complexes. Candidate agents can be small organic compounds having a molecular weight of more than about 50 and less than about 2,500 daltons. Candidate agents can comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and can include at least an amine, carbonyl, hydroxyl or carboxyl group, and can contain at least two of the functional chemical groups. The candidate agents can comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules, including oligonucleotides, polynucleotides, and fragments thereof, depsipeptides, polypeptides and fragments thereof, oligosaccharides, polysaccharides and fragments thereof, lipids, fatty acids, steroids, purines, pyrimidines, derivatives thereof, structural analogs, modified nucleic acids, modified, derivatized or designer amino acids, or combinations thereof.

[0100] An "agent which modulates a biological activity of a subject polypeptide," as used herein, describes any substance, synthetic, semi-synthetic, or natural, organic or inorganic, small molecule or macromolecular, pharmaceutical or protein, with the capability of altering a biological activity of a subject polypeptide or of a fragment thereof, as described herein. Generally, a plurality of assay mixtures is run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, *i.e.*, at zero concentration or below the level of detection. The biological activity can be measured using any assay known in the art.

[0101] An agent which modulates a biological activity of a subject polypeptide increases or decreases the activity at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 50%, at least about 100%, or at least about 2-fold, at least about 5-fold, or at least about 10-fold or more when compared to a suitable control.

[0102] The term "agonist" refers to a substance that mimics the function of an active molecule. Agonists include, but are not limited to, drugs, hormones, antibodies, and neurotransmitters, as well as analogues and fragments thereof.

[0103] The term "antagonist" refers to a molecule that competes for the binding sites of an agonist, but does not induce an active response. Antagonists include, but are not

limited to, drugs, hormones, antibodies, and neurotransmitters, as well as analogues and fragments thereof.

[0104] The term "receptor" refers to a polypeptide that binds to a specific extracellular molecule and may initiate a cellular response.

[0105] The term "ligand" refers to any molecule that binds to a specific site on another molecule.

[0106] The term "modulate" encompasses an increase or a decrease, a stimulation, inhibition, or blockage in the measured activity when compared to a suitable control. "Modulation" of expression levels includes increasing the level and decreasing the level of an mRNA or polypeptide encoded by a polynucleotide of the invention when compared to a control lacking the agent being tested. In some embodiments, agents of particular interest are those which inhibit a biological activity of a subject polypeptide, and/or which reduce a level of a subject polypeptide in a cell, and/or which reduce a level of a subject mRNA in a cell and/or which reduce the release of a subject polypeptide from a eukaryotic cell. In other embodiments, agents of interest are those that increase a biological activity of a subject polypeptide, and/or which increase a level of a subject polypeptide in a cell, and/or which increase a level of a subject mRNA in a cell and/or which increase the release of a subject polypeptide from a eukaryotic cell.

[0107] An agent that "modulates the level of expression of a nucleic acid" in a cell is one that brings about an increase or decrease of at least about 1.25-fold, at least about 1.5-fold, at least about 2-fold, at least about 5-fold, at least about 10-fold, or more in the level (*i.e.*, an amount) of mRNA and/or polypeptide following cell contact with a candidate agent compared to a control lacking the agent.

[0108] "Modulating a level of active subject polypeptide" includes increasing or decreasing activity of a subject polypeptide; increasing or decreasing a level of active polypeptide protein; increasing or decreasing a level of mRNA encoding active subject polypeptide, and increasing or decreasing the release of subject polypeptide for a eukaryotic cell. In some embodiments, an agent is a subject polypeptide, where the subject polypeptide itself is administered to an individual. In some embodiments, an agent is an antibody specific for a subject polypeptide. In some embodiments, an agent is a chemical compound such as a small molecule that may be useful as an orally available drug. Such modulation includes the recruitment of other molecules that directly effect the modulation. For example, an antibody that modulates the activity of a subject polypeptide that is a receptor on a cell surface may

bind to the receptor and fix complement, activating the complement cascade and resulting in lysis of the cell.

[0109] The term "over-expressed" refers to a state wherein there exists any measurable increase over normal or baseline levels. For example, a molecule that is over-expressed in a disorder is one that is manifest in a measurably higher level compared to levels in the absence of the disorder.

[0110] A "stem cell" is a pluripotent or multipotent cell with the ability to self-renew, to remain undifferentiated, and to become differentiated. Stem cells can divide without limit, at least for the lifetime of the animal in which they naturally reside. Stem cells are not terminally differentiated, *i.e.*, they are not at the end of a pathway of differentiation. When a stem cell divides, each daughter cell can either remain a stem cell or it can embark on a course that leads to terminal differentiation.

[0111] An "embryonic stem cell" is a stem cell that is present in or isolated from an embryo. An "adult stem cell" is a stem cell that is present in or isolated from an adult. Either can be pluripotent, having the capacity to differentiate into each and every cell present in the organism, or multipotent, with the ability to differentiate into more than one cell type. Embryonic stem cells derived from the inner cell mass of the embryo can act as pluripotent cells when placed into host blastocysts. Adult stem cells are more frequently multipotent than pluripotent; examples of multipotent adult stem cells include hematopoietic stem cells, peripheral nervous system stem cells, central nervous system stem cells, myogenic stem cells, and mesenchymal stem cells.

[0112] A "mesenchymal stem cell" (MSC) is an adult pluripotent stem cell progenitor of multiple mesenchymal lineages, including bone, cartilage, muscle, fat tissue, marrow stroma, and astrocytes. Mesenchyme is embryonic tissue of mesodermal origin, *i.e.*, tissue that derives from the middle of three germ layers. The mesenchyme is populated by mesenchymal cells, which are typically stellate or fusiform in shape. The embryonic mesoderm gives rise to the musculoskeletal, blood, vascular, and urogenital systems, as well as connective tissue, *i.e.*, the dermis.

[0113] A "hematopoietic" cell is a cell involved in the process of hematopoiesis, *i.e.*, the process of forming mature red and white blood cells from precursor cells. In the adult, hematopoiesis takes place in the bone marrow. Earlier in development, hematopoiesis takes place at different sites during different stages of development; primitive blood cells arise in the yolk sac, and later, blood cells are formed in the liver, spleen, and bone marrow.

Hematopoiesis undergoes complex regulation, including regulation by hormones, *e.g.*, erythropoietin; growth factors, *e.g.*, colony stimulating factors; and cytokines, *e.g.*, interleukins. While the B-lymphocytic component of white blood cells matures in the bone marrow, the T-lymphocytic component of white blood cells matures in the thymus.

[0114] "Differentiation" is a progressive developmental change to a more specialized form or function. Cell differentiation is the process a cell undergoes as it matures to become an overtly specialized cell type. Differentiated cells have distinct characteristics, perform specific functions, and are less likely to divide than their less differentiated counterparts. An "undifferentiated" cell, *e.g.*, an immature, embryonic, or primitive cell, typically has a non-specific appearance, may perform multiple, non-specific activities, and may perform poorly, if at all, in functions typically performed by differentiated cells.

[0115] "Dedifferentiation" is a process by which a mature cell returns to a less mature state. A "dedifferentiated cell" is one that has fewer characteristics of differentiation than it possesses at an earlier point in time. A "dedifferentiated state" is one in which a mature cell has returned or is returning to a less differentiated state, *e.g.*, as in some cancers.

[0116] A "differentiation factor" is a factor that induces a cell to undergo a change in the direction of an overtly specialized cell type. An "anti-differentiation factor" is a factor that prevents or inhibits a cell from undergoing a change in the direction toward an overtly specialized cell type.

[0117] A "co-factor" is a molecule that acts in concert with another substance to bring about certain effects.

[0118] A "lymphokine" is a cytokine produced by a leukocyte, which acts upon another cell. Examples include interleukins, interferon-alpha, tumor necrosis factor-alpha, and granulocyte/monocyte colony-stimulating factor.

[0119] An "anti-inflammatory molecule" is a molecule that can diminish, eliminate, or prevent a response to injury or infection. For example, an antihistamine can counteract the effect of the inflammatory mediator histamine.

[0120] An "anti-cancer molecule" is a molecule that can diminish, eliminate, or prevent the effects of cancer. It includes pharmaceuticals and antibodies.

[0121] An "apoptotic molecule" is a molecule that induces a cell to move towards apoptosis, or programmed cell death. Normally functioning cells undergo apoptosis when their age or their state of health so dictates. Apoptosis is an active process requiring metabolic activity by the dying cell, often characterized by cleavage of the DNA into

fragments. Cells that die by apoptosis do not generally elicit the inflammatory response associated with necrosis. Cancer cells do not typically undergo normal apoptosis.

[0122] First and second therapeutic molecules working in "conjunction" means they work in association with one another to achieve a therapeutic effect.

[0123] First and second heterologous nucleic acid sequences that "interact" with one another means they have an effect on one another such that one of the sequences influences the other. Either may act upon the other, or both may act upon each other.

[0124] A "promoter" is a region of DNA that binds RNA polymerase before initiating the transcription of DNA into RNA. The nucleotide at which transcription begins is designated +1; nucleotides are numbered from this reference point. Negative numbers indicate upstream nucleotides and positive numbers indicate downstream nucleotides. The promoter directs the RNA polymerase to bind to DNA, to open the DNA helix, and to begin RNA synthesis. Some promoters are "constitutive," and direct transcription in the absence of regulatory influences. Some promoters are "tissue specific," and initiate transcription exclusively or selectively in one or a few tissue types. Some promoters are "inducible," and effect gene transcription under the influence of an inducer. Induction can occur, *e.g.*, as the result of a physiologic response, a response to outside signals, or as the result of artificial manipulation.

[0125] A "knockout" mouse is a mouse in which a normal functional gene has been replaced by a non-functional form of the gene, and the function of that particular gene is eliminated. They are typically produced by transplanting embryonic stem cells heterozygous for a knockout mutation in a gene of interest and homozygous for a marker gene, *e.g.*, black coat color into the blastocoel cavity of embryos that are homozygous for an alternate marker, *e.g.*, white coat color. The early embryos then are implanted into a pseudopregnant female. Some of the resulting progeny are chimeras, indicated by the phenotype produced by the marker, *e.g.*, a black and white coat. Chimeric mice are backcrossed to mice with the alternate marker. Progeny from this mating that display the marker present in the mice with the gene of interest (*e.g.*, black coat) have embryonic stem-derived cells in their germ line; DNA analysis of these mice can identify the mice heterozygous for the null allele of the gene of interest, *i.e.*, the "knockout" allele. Intercrossing these heterozygous mice produces mice homozygous for the disrupted allele, *i.e.*, "knockout" mice (Capecchi, 1989).

[0126] Gene "knockout" produces model systems for studying inherited human diseases, investigating the nature of genetic diseases and the efficacy of different types of

treatment, and for developing effective gene therapies to cure these diseases. For example, a "knockout" line of mutant mice homozygous for a null allele of the cystic fibrosis transmembrane regulator gene, demonstrates symptoms similar to those of humans with cystic fibrosis. These mice provide a model system for studying this genetic disease and developing effective therapies.

[0127] A "transgenic mouse" is a mouse that has stably incorporated one or more genes from another cell or organism and can pass them on to successive generations. Transgenic mice with an exogenous DNA sequence of interest integrated into its DNA are typically produced by injecting DNA containing a gene of interest into one of the two pronuclei (the male and female haploid nuclei contributed by the parents) of a fertilized mouse egg before they fuse. The injected DNA is randomly integrated into the chromosomes of the diploid zygote. Injected eggs then are transferred to foster mothers in which normal cell growth and differentiation occurs. Some of the progeny will contain the exogenous DNA, and breeding and backcrossing can produce pure transgenic strains homozygous for the transgene (Brinster et al., 1981).

[0128] Transgenic mice are useful for studying various aspects of normal mammalian biology, and also provide a model system for studying disease processes. For example, many forms of cancer are promoted by normal cellular *myc* genes acting in a dominant fashion owing to their misregulated activity. Transgenic mice carrying the *myc* gene develop normally, and form tumors at a high frequency in a subset of cells that express the transgene.

[0129] A "therapeutic factor" encoded by a first heterologous nucleic acid sequence of a modified mesenchymal cell is a factor, excluding a cell survival factor (Mangi et al., 2003; WO 03/073998), that is preventative, palliative, curative, or otherwise useful in treating or ameliorating, or preventing the recurrence of a disease, disorder, syndrome or condition, and is not an anti-cancer agent.

[0130] "Telomerase" is a DNA polymerase enzyme that selectively elongates DNA from the telomere, *i.e.*, the end of a chromosome. Telomeric DNA contains multiple, *e.g.*, hundreds, of tandem repeats of a hexanucleotide sequence. One strand of telomeric DNA is G-rich at the 3' end, and slightly longer than the other strand. Telomeric DNA can form large duplex loops, wherein the single-stranded region at the very end of the structure loops back to form a DNA duplex with another part of the repeated sequence, displacing a part of the original telomeric duplex. This looplike structure is formed and stabilized by specific telomere-binding proteins. These structures protect and mask the end of the chromosome.

[0131] The telomeric looplike structures are generated by telomerase. The telomerase enzyme contains an RNA molecule that serves as the template for elongating the G-rich strand of telomeric DNA. Thus, the enzyme carries the information necessary to generate the telomere sequences. Telomerases also have a protein component, which is related to reverse transcriptases. Telomerases can influence cell aging, and play a role in cellular cancer biology.

[0132] "Tumor necrosis factor" (TNF) encompasses a family of receptor ligands that display pleiotropic effects on normal and malignant cells. Natural induction of TNF is protective, but its overproduction may be detrimental and even lethal to the host. TNF elicits a variety of responses in different cell types. TNF was originally characterized as an antitumor agent and a cytotoxic factor for malignant cells. It subverts the electron transport system of mitochondria to produce oxygen radicals, which can kill malignant cells lacking protective enzymes. TNF also plays a role in the defense against viral, bacterial, and parasitic infections, and in mediating autoimmune responses (Fiers, 1991). TNF inhibitors have been used to treat psoriasis (Weinberg and Saini, 2003).

[0133] "Treatment," "treating," and the like, as used herein, refer to obtaining a desired pharmacologic and/or physiologic effect, covering any treatment of a pathological condition or disorder in a mammal, including a human. The effect may be prophylactic in terms of completely or partially preventing a disorder or symptom thereof and/or may be therapeutic in terms of a partial or complete cure for a disorder and/or adverse affect attributable to the disorder. That is, "treatment" includes (1) preventing the disorder from occurring or recurring in a subject who may be predisposed to the disorder but has not yet been diagnosed as having it, (2) inhibiting the disorder, such as arresting its development, (3) stopping or terminating the disorder or at least symptoms associated therewith, so that the host no longer suffers from the disorder or its symptoms, such as causing regression of the disorder or its symptoms, for example, by restoring or repairing a lost, missing or defective function, or stimulating an inefficient process, or (4) relieving, alleviating, or ameliorating the disorder, or symptoms associated therewith, where ameliorating is used in a broad sense to refer to at least a reduction in the magnitude of a parameter, such as inflammation, pain, and/or tumor size.

[0134] A "pharmaceutically acceptable carrier," "pharmaceutically acceptable diluent," or "pharmaceutically acceptable excipient," or "pharmaceutically acceptable vehicle," used interchangeably herein, refer to a non-toxic solid, semisolid or liquid filler,

diluent, encapsulating material or formulation auxiliary of any conventional type. A pharmaceutically acceptable carrier is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the carrier for a formulation containing polypeptides would not normally include oxidizing agents and other compounds that are known to be deleterious to polypeptides. Suitable carriers include, but are not limited to, water, dextrose, glycerol, saline, ethanol, and combinations thereof. The carrier can contain additional agents such as wetting or emulsifying agents, pH buffering agents, or adjuvants which enhance the effectiveness of the formulation. Adjuvants of the invention include, but are not limited to Freund's, Montanide ISA Adjuvants [Seppic, Paris, France], Ribi's Adjuvants (Ribi ImmunoChem Research, Inc., Hamilton, MT), Hunter's TiterMax (CytRx Corp., Norcross, GA), Aluminum Salt Adjuvants (Alhydrogel - Superfos of Denmark/Accurate Chemical and Scientific Co., Westbury, NY), Nitrocellulose-Adsorbed Protein, Encapsulated Antigens, and Gerbu Adjuvant (Gerbu Biotechnik GmbH, Gaiberg, Germany/C-C Biotech, Poway, CA). Topical carriers include liquid petroleum, isopropyl palmitate, polyethylene glycol, ethanol (95%), polyoxyethylene monolaurate (5%) in water, or sodium lauryl sulfate (5%) in water. Other materials such as anti-oxidants, humectants, viscosity stabilizers, and similar agents can be added as necessary. Percutaneous penetration enhancers such as Azone can also be included.

[0135] "Pharmaceutically acceptable salts" include the acid addition salts (formed with the free amino groups of the polypeptide) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, mandelic, oxalic, and tartaric. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, and histidine.

[0136] Compositions for oral administration can form solutions, suspensions, tablets, pills, capsules, sustained release formulations, oral rinses, or powders.

[0137] The term "unit dosage form," as used herein, refers to physically discrete units suitable as unitary dosages for human and animal subjects, each unit containing a predetermined quantity of compounds of the present invention calculated in an "effective amount," that is, a dosage sufficient to produce the desired result or effect in association with a pharmaceutically acceptable carrier. The specifications for the novel unit dosage forms of

the present invention depend on the particular compound employed, the host, and the effect to be achieved, as well as the pharmacodynamics associated with each compound in the host.

Compositions

[0138] The present invention provides novel isolated polynucleotides encoding polypeptides and fragments thereof. The present invention also provides novel isolated polypeptides, fragments thereof, and compositions comprising same. The present invention further provides polynucleotide compositions that can be used to identify the polypeptides.

[0139] The present invention provides recombinant vectors and host cells for use in gene expression, primer pairs for use in hybridizations, computer-based embodiments for use in bioinformatics, and transgenic animals and embryonic stem cell lines for use in mutating and regulating gene expression.

Nucleic Acids

Sequences

[0140] This invention provides genes encoding proteins, the encoded proteins, and fragments and homologs thereof. It provides human polynucleotide sequences and the corresponding mouse polynucleotide sequences.

[0141] The nucleic acids of the subject invention can encode all or a part of the subject proteins. Double or single stranded fragments can be obtained from the DNA sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, for example by restriction enzyme digestion or polymerase chain reaction (PCR) amplification. The use of the polymerase chain reaction has been described (Saiki et al., 1985) and current techniques have been reviewed (Sambrook et al., 1989; McPherson et al. 2000; Dieffenbach and Dveksler, 1995). For the most part, DNA fragments will be of at least about 5 nucleotides, at least about 8 nucleotides, at least about 10 nucleotides, at least about 15 nucleotides, at least about 18 nucleotides, at least about 20 nucleotides, at least about 25 nucleotides, at least about 30 nucleotides, or at least about 50 nucleotides, at least about 75 nucleotides, or at least about 100 nucleotides. Nucleic acid compositions that encode at least six contiguous amino acids (*i.e.*, fragments of 18 nucleotides or more), for example, nucleic acid compositions encoding at least 8 contiguous amino acids (*i.e.*, fragments of 24 nucleotides or more), are useful in directing the expression or the synthesis of peptides that can be used as immunogens (Lerner, 1982; Shinnick et al., 1983; Sutcliffe et al., 1983).

[0142] In some embodiments, a polynucleotide of the invention comprises a nucleotide sequence of at least about 5, at least about 8, at least about 10, at least about 15, at

least about 18, at least about 20, at least about 25, at least about 30, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, at least about 1000, at least about 1100, at least about 1200, at least about 1300, at least about 1400, at least about 1500, at least about 1600, at least about 1700, at least about 1800, at least about 1900, at least about 2000, at least about 2100, at least about 2200, at least about 2300, at least about 2400, at least about 2500, at least about 3000, at least about 4000, or at least about 5000 contiguous nucleotides of any one of the sequences shown in SEQ. ID. NOS. 1-187 and 375-484, or the coding region thereof, or a complement thereof.

[0143] In other embodiments, a polynucleotide of the invention has at least about 60%, 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97%, at least about 98%, or at least about 99% nucleotide sequence identity with a nucleotide sequence, or a fragment thereof, of the coding region of any one of the sequences shown in SEQ. ID. NOS. 1-187 and 375-484, or a complement thereof. These sequence variants include naturally-occurring variants (*e.g.*, SNPs, allelic variants, and homologs from other species), degenerate variants, variants associated with disease or pathological states, and variants resulting from random or directed mutagenesis, as well as from chemical or other modification.

[0144] In some embodiments, a polynucleotide of the invention comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence of at least about 5, at least about 8, at least about 10, at least about 15, at least about 18, at least about 20, at least about 25, at least about 30, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, or at least about 1000 contiguous amino acids of at least one of the sequences shown in SEQ. ID. NOS. 188-374 (*e.g.*, a polypeptide encoded by at least one of the nucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484), up to and including an entire amino acid sequence as shown in SEQ. ID. NOS. 188-374 (or as encoded by at least one of the nucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484).

[0145] In some embodiment, the present invention includes the present polynucleotide selected from SEQ. ID. NOS. 1-187 and 375-484, which contain 300 bp of 5' terminus of a protein encoding polynucleotide sequence. Such a polynucleotide is useful for the purposes of clustering gene sequences to determine gene family.

[0146] In further embodiments, a polynucleotide of the invention hybridizes under stringent hybridization conditions to a polynucleotide having the coding region of any one of the sequences shown in SEQ. ID. NOS. 1-187 and 375-484, or a complement thereof.

[0147] The polynucleotides of the invention include those that encode variants of the polypeptide sequences encoded by the polynucleotides of the appendices. In some embodiments, these polynucleotides encode variant polypeptides that include insertions, additions, deletions, or substitutions compared with the polypeptides encoded by the nucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484, and in Table 1. Conservative amino acid substitutions include serine/threonine, valine/leucine/isoleucine, asparagine/histidine/glutamine, glutamic acid/aspartic acid, etc. (Gonnet et al., 1992).

[0148] The nucleic acids of the invention include degenerate variants that can be translated, according to the standard genetic code, to provide an amino acid sequence identical to that translated from the nucleic acid sequences herein. For example, synonymous codons include GGG, GGA, GGC, and GGU, each encoding Glycine.

[0149] The nucleic acids of the invention include single nucleotide polymorphisms (SNPs), which occur frequently in eukaryotic genomes (Lander, et al. 2001). The nucleotide sequence determined from one individual of a species can differ from other allelic forms present within the population.

[0150] The nucleic acids of the invention include homologs of the polynucleotides. The source of homologous genes can be any species, *e.g.*, primate species, particularly human; rodents, such as rats, hamsters, guinea pigs, and mice; rabbits, canines, felines; cattle, such as bovines; goats, pigs, sheep, equines, crustaceans, birds, chickens, reptiles, amphibians, fish, insects, plants, fungi, yeast, nematodes, etc. Among mammalian species, *e.g.*, human and mouse, homologs have substantial sequence similarity, *e.g.*, at least about 60% sequence identity, at least about 75% sequence identity, or at least about 80% sequence identity among nucleotide sequences. In many embodiments of interest, homology will be at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97%, or at least about 98%, where in certain embodiments of interest homology will be as high as about 99%.

[0151] Modifications in the native structure of nucleic acids, including alterations in the backbone, sugars or heterocyclic bases, have been shown to increase intracellular stability and binding affinity. Among useful changes in the backbone chemistry are phosphorothioates; phosphorodithioates, where both of the non-bridging oxygens are substituted with sulfur; phosphoroamidites; alkyl phosphotriesters and boranophosphates. Achiral phosphate derivatives include 3'-O'-5'-S-phosphorothioate, 3'-S-5'-O-phosphorothioate, 3'-CH₂-5'-O-phosphonate and 3'-NH-5'-O-phosphoroamidate. Peptide nucleic acids replace the entire ribose phosphodiester backbone with a peptide linkage.

[0152] Sugar modifications are also used to enhance stability and affinity. The α -anomer of deoxyribose can be used, where the base is inverted with respect to the natural β -anomer. The 2'-OH of the ribose sugar can be altered to form 2'-O- methyl or 2'-O-allyl sugars, which provides resistance to degradation without comprising affinity.

[0153] Modification of the heterocyclic bases must maintain proper base pairing. Some useful substitutions include deoxyuridine for deoxythymidine; 5-methyl-2'-deoxycytidine and 5-bromo-2'-deoxycytidine for deoxycytidine. 5- propynyl-2'-deoxyuridine and 5-propynyl-2'-deoxycytidine have been shown to increase affinity and biological activity when substituted for deoxythymidine and deoxycytidine, respectively.

[0154] A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. It can further include the 3' and 5' untranslated regions found in the mature mRNA. It can further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, etc., including about 1 kb, about 2 kb, and possibly more, of flanking genomic DNA at either the 5' or 3' end of the transcribed region. The genomic DNA can be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' or 5', or internal regulatory sequences as sometimes found in introns, contains sequences required for proper tissue and stage specific expression.

[0155] Nucleic acid molecules of the invention can comprise heterologous nucleic acid molecules, *i.e.*, nucleic acid molecules other than the subject nucleic acid molecules, of any length. For example, the subject nucleic acid molecules can be flanked on the 5' and/or 3' ends by heterologous nucleic acid molecules of from about 1 nucleotide to about 10 nucleotides, from about 10 nucleotides to about 20 nucleotides, from about 20 nucleotides to about 50 nucleotides, from about 50 nucleotides to about 100 nucleotides, from about 100

nucleotides to about 250 nucleotides, from about 250 nucleotides to about 500 nucleotides, or from about 500 nucleotides to about 1000 nucleotides, or more in length.

[0156] The subject polynucleotides include those that encode fusion proteins comprising the subject polypeptides fused to "fusion partners." For example, the present soluble receptor or ligand can be fused to an immunoglobulin fragment, such as an Fc fragment for stability in circulation or to fix complement. Other polypeptide fragments that have equivalent capabilities as the Fc fragments can also be used herein.

[0157] The isolated nucleic acids of the invention can be used as probes to detect and characterize gross alteration in a genomic locus, such as deletions, insertions, translocations, and duplications, *e.g.*, applying fluorescence *in situ* hybridization (FISH) techniques to examine chromosome spreads (Andreeff et al., 1999). The nucleic acids are also useful for detecting smaller genomic alterations, such as deletions, insertions, additions, translocations, and substitutions (*e.g.*, SNPs).

[0158] When used as probes to detect nucleic acid molecules capable of hybridizing with nucleic acids described in the appendices, the nucleic acid molecules can be flanked by heterologous sequences of any length. When used as probes, a subject nucleic acid can include nucleotide analogs that incorporate labels that are directly detectable, such as radiolabels or fluorophores, or nucleotide analogs that incorporate labels that can be visualized in a subsequent reaction, such as biotin or various haptens. Haptens that are commonly conjugated to nucleotides for subsequent labeling include biotin, digoxigenin, and dinitrophenyl.

[0159] Suitable fluorescent labels include fluorochromes *e.g.*, fluorescein and its derivatives, *e.g.*, fluorescein isothiocyanate (FITC6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE),), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM); coumarin and its derivatives, *e.g.*, 7-amino-4-methylcoumarin, aminocoumarin; bodipy dyes, such as Bodipy FL; cascade blue; Oregon green; rhodamine dyes, *e.g.*, rhodamine, 6-carboxy-X-rhodamine (ROX), Texas red, phycoerythrin, and tetramethylrhodamine; eosins and erythrosins; cyanine dyes, *e.g.*, allophycocyanin, Cy3 and Cy5 or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); macrocyclic chelates of lanthanide ions, *e.g.*, quantum dye, etc; and chemiluminescent molecules, *e.g.*, luciferases.

[0160] Fluorescent labels also include a green fluorescent protein (GFP), *i.e.*, a "humanized" version of a GFP, *e.g.*, wherein codons of the naturally-occurring nucleotide

sequence are changed to more closely match human codon bias; a GFP derived from *Aequoria victoria* or a derivative thereof, *e.g.*, a "humanized" derivative such as Enhanced GFP, which are available commercially, *e.g.*, from Clontech, Inc.; other fluorescent mutants of a GFP from *Aequoria victoria*, *e.g.*, as described in U.S. Patent No. 6,066,476; 6,020,192; 5,985,577; 5,976,796; 5,968,750; 5,968,738; 5,958,713; 5,919,445; 5,874,304; a GFP from another species such as *Renilla reniformis*, *Renilla mulleri*, or *Ptilosarcus guernyi*, as previously described (WO 99/49019; Peelle et al., 2001), "humanized" recombinant GFP (hrGFP) (Stratagene[®]); any of a variety of fluorescent and colored proteins from Anthozoan species, (*e.g.*, Matz et al., 1999).

[0161] Probes can also contain fluorescent analogs, including commercially available fluorescent nucleotide analogs that can readily be incorporated into a subject nucleic acid. These include deoxyribonucleotides and/or ribonucleotide analogs labeled with Cy3, Cy5, Texas Red, Alexa Fluor dyes, rhodamine, cascade blue, or BODIPY, and the like.

[0162] Suitable radioactive labels include, *e.g.*, ³²P, ³⁵S, or ³H. For example, probes can contain radiolabeled analogs, including those commonly labeled with ³²P or ³⁵S, such as α -³²P-dATP, -dTTP, -dCTP, and dGTP; γ -³⁵S-GTP and α -³⁵S-dATP, and the like.

[0163] Nucleic acids of the invention can also be bound to a substrate. Subject nucleic acids can be attached covalently, attached to a surface of the support or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence, *e.g.*, by noncovalent interactions, or some combination thereof. The nucleic acids can be bound to a substrate to which a plurality of other nucleic acids are concurrently bound, hybridization to each of the plurality of the bound nucleic acids being separately detectable.

[0164] The substrate can be porous or solid, planar or non-planar, unitary or distributed; and the bond between the nucleic acid and the substrate can be covalent or non-covalent. The substrate can be in the form of microbeads or nanobeads. Substrates include, but are not limited to, a membrane, such as nitrocellulose, nylon, positively-charged derivatized nylon; a solid substrate such as glass, amorphous silicon, crystalline silicon, plastics (including *e.g.*, polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, cellulose acetate, or mixtures thereof).

[0165] The subject nucleic acids include antisense RNA, ribozymes, and RNAi. Further, The nucleic acids of the invention can be used for antisense or RNAi inhibition of

transcription or translation using methods known in the art (Phillips, 1999a; Phillips, 1999b; Hartmann et al., 1999; Stein et al., 1998; Agrawal et al., 1998).

Expression Vectors

[0166] The instant invention further provides host cells, *e.g.*, recombinant host cells, that comprise a subject nucleic acid, host cells that comprise a recombinant vector, and host cells that secrete antibodies of the invention. Subject host cells can be cultured *in vitro*, or can be part of a multicellular organism. Host cells are described in more detail below. The instant invention further provides transgenic plants and non-human animals, as described in more detail below.

[0167] In addition to the plurality of uses described in greater detail in following sections, the subject nucleic acids find use in the preparation of all or a portion of the polypeptides of the subject invention, as described above, using an expression system. For expression, an expression vector can be employed. The expression vector will provide a transcriptional and translational initiation region, which may be inducible, conditionally-active, or constitutive, or tissue-specific, where the coding region is operably linked under the transcriptional control of the transcriptional initiation region, and a transcriptional and translational termination region. These control regions can be native to a gene encoding the subject peptides, or can be derived from heterologous or exogenous sources.

[0168] The subject nucleic acids can also be provided as part of a vector (*e.g.*, a polynucleotide construct comprising an expression cassette), a wide variety of which are known in the art. Vectors include, but are not limited to, plasmids; cosmids; viral vectors; human, yeast, bacterial, P1-derived artificial chromosomes (HAC's, YAC's, BAC's, PAC's, etc.), mini-chromosomes, and the like. Vectors are amply described in numerous publications well known to those in the art (Ausubel, et al.; Jones et al., 1998a; Jones et al., 1998b). Vectors can provide for nucleic acid expression, for nucleic acid propagation, or both.

[0169] Vectors typically include at least one origin of replication, at least one site for insertion of heterologous nucleic acid (*e.g.*, in the form of a polylinker with multiple, tightly clustered, single cutting restriction endonuclease recognition sites), and at least one selectable marker, although some integrative vectors will lack an origin that is functional in the host to be chromosomally modified, and some vectors will lack selectable markers. Vectors are transiently or stably maintained in the cells, usually for a period of at least about one day, at least about several days to at least about several weeks.

[0170] Promoters of the invention can be naturally contiguous or not naturally contiguous to the expressed nucleic acid molecule. The promoters can be inducible, conditionally active (such as the cre-lox promoter), constitutive, and/or tissue specific.

[0171] Prior to vector insertion, the DNA of interest will be obtained substantially free of other nucleic acid sequences. The DNA can be "recombinant," and flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

[0172] Expression vectors generally have convenient restriction sites located near the promoter sequence to provide for the insertion of nucleic acid sequences encoding heterologous protein or RNA molecules. A selectable marker operative in the expression system or host can be present. Expression vectors can be used for the production of fusion proteins, where the fusion peptide provides additional functionality, *i.e.*, increased protein synthesis, a leader sequence for secretion, stability, reactivity with defined antisera, or an enzyme marker, *e.g.*, β -galactosidase.

[0173] Expression vectors can be prepared comprising a transcription cassette comprising a transcription initiation region, the gene or fragment thereof, and a transcriptional termination region. Of particular interest is the use of DNA sequences that allow for the expression of functional epitopes or domains, at least about 5, at least about 8, at least about 10, at least about 15, at least about 18, at least about 20, at least about 25, at least about 30, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, or at least about 1000 amino acids in length, or any of the above-described fragments, up to and including the complete open reading frame of the gene. After introduction of these DNA sequences, the cells containing the vector construct can be selected by means of a selectable marker, and the selected cells expanded and used as expression-competent host cells.

[0174] Host cells can comprise prokaryotes or eukaryotes that express proteins and polypeptides in accordance with conventional methods, the method depending on the purpose for expression. For large scale production of the protein, a unicellular organism, such as *E. coli*, *B. subtilis*, *S. cerevisiae*, insect cells in combination with baculovirus vectors, or cells of a higher organism such as vertebrates, particularly mammals, *e.g.*, COS 7 cells, can be used as

the expression host cells. In some situations, it is desirable to express eukaryotic genes in eukaryotic cells, where the encoded protein will benefit from native folding and post-translational modifications.

[0175] Specific expression systems of interest include plants, bacteria, yeast, insect cells, and mammalian cell-derived expression systems. Representative systems from each of these categories are provided below.

[0176] Expression systems in plants include those described in U.S. Patent No. 6,096,546 and U.S. Patent No. 6,127,145.

[0177] Expression systems in bacteria include those described by Chang et al., 1978; Goeddel et al., 1979; Goeddel et al., 1980; EP 0 036,776; U.S. Patent No. 4,551,433; DeBoer et al., 1983); and Siebenlist et al., 1980.

[0178] Expression systems in yeast include those described by Hinnen et al., 1978; Ito et al., 1983; Kurtz et al., 1986; Kunze et al., 1985; Gleeson et al., 1986; Roggenkamp et al., 1986; Das et al., 1984; De Louvencourt et al., 1983; Van den Berg et al., 1990; Kunze et al., 1985; Cregg et al., 1985; U.S. Patent Nos. 4,837,148 and 4,929,555; Beach and Nurse, 1981; Davidow et al., 1985; Gaillardin et al., 1985; Ballance et al., 1983; Tilburn et al., 1983; Yelton et al., 1984; Kelly and Hynes, 1985; EP 0 244,234; WO 91/00357; and U.S. Patent No. 6,080,559.

[0179] Expression systems for heterologous genes in insects include those described in U.S. Patent No. 4,745,051; Friesen et al., 1986; EP 0 127,839; EP 0 155,476; Vlak et al., 1988; Miller et al., 1988; Carbonell et al., 1988; Maeda et al., 1985; Lebacqz-Verheyden et al., 1988; Smith et al., 1985); Miyajima et al., 1987; and Martin et al., 1988. Numerous baculoviral strains and variants and corresponding permissive insect host cells are described in Luckow et al., 1988, Miller et al., 1986, and Maeda et al., 1985. The insect cell expression system is useful not only for production of heterologous proteins intracellularly, but can be used for expression of transmembrane proteins on the insect cell surfaces. Such insect cells can be used as immunogen for production of antibodies, for example, by injection of the insect cells into mice or rabbits or other suitable animals, for production of antibodies.

[0180] Mammalian expression systems include those described in Dijkema et al., 1985; Gorman et al., 1982; Boshart et al., 1985; and U.S. Patent No. 4,399,216. Additional features of mammalian expression are facilitated as described in Ham and Wallace, 1979; Barnes and Sato, 1980 U.S. Patent Nos. 4,767,704, 4,657,866, 4,927,762, 4,560,655, WO

90/103430, WO 87/00195, and U.S. RE 30,985. Mammalian cell expression systems can also be used for production of antibodies.

[0181] The present polynucleotides can also be used in cell-free expression systems such as bacterial system, *e.g.*, *E. coli* lysate, rabbit reticulocyte lysate system, wheat germ extract system, frog oocyte lysate system, and the like which is conventional in the art. See, for example, WO 00/68412, WO 01/27260, WO 02/24939, WO 02/38790, WO 91/02076, and WO 91/02075.

[0182] When any of the above-referenced host cells, or other appropriate host cells or organisms, are used to replicate and/or express the polynucleotides of the invention, the resulting replicated nucleic acid, RNA, expressed protein or polypeptide, is within the scope of the invention as a product of the host cell or organism.

[0183] Once the gene corresponding to a selected polynucleotide is identified, its expression can be regulated in the gene's native cell types. For example, an endogenous gene of a cell can be regulated by an exogenous regulatory sequence inserted into the genome of the cell at a location that will enhance or reduce expression of the gene corresponding to the subject polypeptide. The regulatory sequence can be designed to integrate into the genome via homologous recombination, as disclosed in U.S. Patent Nos. 5,641,670 and 5,733,761, the disclosures of which are herein incorporated by reference. Alternatively, it can be designed to integrate into the genome via non-homologous recombination, as described in WO 99/15650, the disclosure of which is also herein incorporated by reference. Also encompassed in the subject invention is the production of proteins without manipulating the encoding nucleic acid itself, but rather by integrating a regulatory sequence into the genome of a cell that already includes a gene that encodes the protein of interest; this production method is described in the above-incorporated patent documents.

Isolated Primer Pairs

[0184] In some embodiments, the invention provides isolated nucleic acids that, when used as primers in a polymerase chain reaction, amplify a subject polynucleotide, or a polynucleotide containing a subject polynucleotide. The amplified polynucleotide is from about 20 to about 50, from about 50 to about 75, from about 75 to about 100, from about 100 to about 125, from about 125 to about 150, from about 150 to about 175, from about 175 to about 200, from about 200 to about 250, from about 250 to about 300, from about 300 to about 350, from about 350 to about 400, from about 400 to about 500, from about 500 to about 600, from about 600 to about 700, from about 700 to about 800, from about 800 to

about 900, from about 900 to about 1000, from about 1000 to about 2000, from about 2000 to about 3000, from about 3000 to about 4000, from about 4000 to about 5000, or from about 5000 to about 6000 nucleotides or more in length.

[0185] The isolated nucleic acids themselves are from about 10 to about 20, from about 20 to about 30, from about 30 to about 40, from about 40 to about 50, from about 50 to about 100, or from about 100 to about 200 nucleotides in length. Generally, the nucleic acids are used in pairs in a polymerase chain reaction, where they are referred to as "forward" and "reverse" primers.

[0186] Thus, in some embodiments, the invention provides a pair of isolated nucleic acid molecules, each from about 10 to about 200 nucleotides in length, the first nucleic acid molecule of the pair comprising a sequence of at least 10 contiguous nucleotides having 100% sequence identity to a nucleic acid sequence as shown in SEQ. ID. NOS. 1-187 and 375-484 and the second nucleic acid molecule of the pair comprising a sequence of at least 10 contiguous nucleotides having 100% sequence identity to the reverse complement of the nucleic acid sequence shown in SEQ. ID. NOS. 1-187 and 375-484, wherein the sequence of the second nucleic acid molecule is located 3' of the nucleic acid sequence of the first nucleic acid molecule shown in SEQ. ID. NOS. 1-187 and 375-484. The primer nucleic acids are prepared using any known method, *e.g.*, automated synthesis, and can be chosen to specifically amplify a cDNA copy of an mRNA encoding a subject polypeptide.

[0187] In some embodiments, the first and/or the second nucleic acid molecules comprise a detectable label. The label can be a radioactive molecule, fluorescent molecule or another molecule, *e.g.*, hapten, as described in detail above. Further, the label can be a two stage system, where the amplified DNA is conjugated to another molecule, *i.e.*, biotin, digoxin, or a hapten, that has a high affinity binding partner, *i.e.*, avidin, antidigoxin, or a specific antibody, respectively, and the binding partner conjugated to a detectable label. The label can be conjugated to one or both of the primers. Alternatively, the pool of nucleotides used in the amplification is labeled, so as to incorporate the label into the amplification product.

[0188] Conditions that increase stringency of both DNA/DNA and DNA/RNA hybridization reactions are widely known and published in the art. See, for example, Sambrook, 1989, and examples provided above. Examples of relevant conditions include (in order of increasing stringency): incubation temperatures of 25°C, 37°C, 50°C, and 68°C; buffer concentrations of 10 x SSC, 6 x SSC, 1 x SSC, 0.1 x SSC (where 1 x SSC is 0.15 M

NaCl and 15 mM citrate buffer); and their equivalents using other buffer systems; formamide concentrations of 0%, 25%, 50%, and 75%; incubation times from 5 minutes to 24 hours; 1, 2, or more washing steps; wash incubation times of 1, 2, or 15 minutes; and wash solutions of 6 x SSC, 1 x SSC, 0.1 x SSC, or deionized water.

[0189] For example, "high stringency conditions" include hybridization in 50% formamide, 5X SSC, 0.2 µg/µl poly(dA), 0.2 µg/µl human cot1 DNA, and 0.5% SDS, in a humid oven at 42°C overnight, followed by successive washes in 1X SSC, 0.2% SDS at 55°C for 5 minutes, followed by washing at 0.1X SSC, 0.2% SDS at 55°C for 20 minutes. Further examples of high stringency conditions include hybridization at 50°C and 0.1×SSC (15 mM sodium chloride/1.5 mM sodium citrate); overnight incubation at 42°C in a solution containing 50% formamide, 1 × SSC (150 mM NaCl, 15 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5 × Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1 × SSC at about 65°. High stringency conditions also include aqueous hybridization (*e.g.*, free of formamide) in 6X SSC (where 20X SSC contains 3.0 M NaCl and 0.3 M sodium citrate), 1% sodium dodecyl sulfate (SDS) at 65°C for about 8 hours (or more), followed by one or more washes in 0.2 X SSC, 0.1% SDS at 65°C. Highly stringent hybridization conditions are hybridization conditions that are at least as stringent as any one of the above representative conditions. Other stringent hybridization conditions are known in the art and can also be employed to identify nucleic acids of this particular embodiment of the invention.

[0190] Conditions of "reduced stringency," suitable for hybridization to molecules encoding structurally and functionally related proteins, or otherwise serving related or associated functions, are the same as those for high stringency conditions but with a reduction in temperature for hybridization and washing to lower temperatures (*e.g.*, room temperature or about 22°C to 25°C). For example, moderate stringency conditions include aqueous hybridization (*e.g.*, free of formamide) in 6X SSC, 1% SDS at 65°C for about 8 hours (or more), followed by one or more washes in 2X SSC, 0.1% SDS at room temperature. Low stringency conditions include, for example, aqueous hybridization at 50°C and 6×SSC (0.9 M sodium chloride/0.09 M sodium citrate) and washing at 25°C in 1×SSC (0.15 M sodium chloride/0.015 M sodium citrate).

[0191] The specificity of a hybridization reaction allows any single-stranded sequence of nucleotides to be labeled with a radioisotope or chemical and used as a probe to find a complementary strand, even in a cell or cell extract that contains millions of different

DNA and RNA sequences. Probes of this type are widely used to detect the nucleic acids corresponding to specific genes, both to facilitate the purification and characterization of the genes after cell lysis and to localize them in cells, tissues, and organisms.

[0192] Moreover, by carrying out hybridization reactions under conditions of "reduced stringency," a probe prepared from one gene can be used to find homologous evolutionary relatives - both in the same organism, where the relatives form part of a gene family, and in other organisms, where the evolutionary history of the nucleotide sequence can be traced. A person skilled in the art would recognize how to modify the conditions to achieve the requisite degree of stringency for a particular hybridization.

Libraries

[0193] The polynucleotide libraries of the invention generally comprise a collection of sequence information of a plurality of polynucleotide sequences, where at least one of the polynucleotides has a sequence shown in SEQ. ID. NOS. 1-187 and 375-484. By plurality is meant at least 2, at least 3, or at least all of the sequences in the appendices. The information may be provided in either biochemical form (*e.g.*, as a collection of polynucleotide molecules), or in electronic form (*e.g.*, as a collection of polynucleotide sequences stored in a computer-readable form, as in a computer-based system, a computer data file, and/or as a part of a computer program). The length and number of polynucleotides in the library will vary with the nature of the library, *e.g.*, if the library is an oligonucleotide array, a cDNA array, or a computer database of the sequence information.

[0194] The sequence information contained in either a biochemical or an electronic library of polynucleotides can be used in a variety of ways, *e.g.*, as a resource for gene discovery, as a representation of sequences expressed in a selected cell type (*e.g.*, cell type markers), or as markers of a given disorder or disease state. In general, a disease marker is a representation of a gene product that is present in all cells affected by disease either at an increased or decreased level relative to a normal cell (*e.g.*, a cell of the same or similar type that is not substantially affected by disease). For example, a polynucleotide sequence in a library can be a polynucleotide that represents an mRNA, polypeptide, or other gene product encoded by the polynucleotide, that is either over-expressed or under-expressed in one cell compared to another (*e.g.*, a first cell type compared to a second cell type; a normal cell compared to a diseased cell; a cell not exposed to a signal or stimulus compared to a cell exposed to that signal or stimulus; and the like).

[0195] The nucleotide sequence information of the library can be embodied in any suitable form, *e.g.*, electronic or biochemical forms. For example, a library of sequence information embodied in electronic form comprises an accessible computer data file that may contain the representative nucleotide sequences of genes that are differentially expressed (*e.g.*, over-expressed or under-expressed) as between, *e.g.*, a first cell type compared to a second cell type (*e.g.*, expression in a brain cell compared to expression in a kidney cell); a normal cell compared to a diseased cell (*e.g.*, a non-cancerous cell compared to a cancerous cell); a cell not exposed to an internal or external signal or stimulus compared to a cell exposed to that signal or stimulus (*e.g.*, a cell contacted with a ligand compared to a control cell not contacted with the ligand); and the like. Other combinations and comparisons of cells will be readily apparent to the ordinarily skilled artisan. Biochemical embodiments of the library include a collection of nucleic acid molecules that have the sequences of the genes in the library, where the nucleic acids can correspond to the entire gene in the library or to a fragment thereof, as described in greater detail below.

[0196] Where the library is an electronic library, the nucleic acid sequence information can be present in a variety of media. For example, the nucleic acid sequences of any of the polynucleotides shown in SEQ. ID. NOS. 1-187 and 375-484 can be recorded on computer readable media of a computer-based system, *e.g.*, any medium that can be read and accessed directly by a computer. One of skill in the art can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising a recording of the present sequence information. Any convenient data storage structure can be chosen, based on the means used to access the stored information. A variety of data processor programs and formats can be used for storage, *e.g.*, word processing text file, database format, etc. In addition to the sequence information, electronic versions of the libraries of the invention can be provided in conjunction or connection with other computer-readable information and/or other types of computer-based files (*e.g.*, searchable files, executable files, etc, including, but not limited to, for example, search program software, etc.).

[0197] By providing the nucleotide sequence in computer readable form in a computer-based system, the information can be accessed for a variety of purposes. Computer software to access sequence information is publicly available. Conventional bioinformatics tools can be utilized to analyze sequences to determine sequence identity, sequence similarity, and gap information. For example, the gapped BLAST (Altschul et al., 1990,

Altschul et al., 1997), and BLAZE (Brutlag et al., 1993) search algorithms on a Sybase system, or the TeraBLAST (TimeLogic, Crystal Bay, Nevada) program optionally running on a specialized computer platform available from TimeLogic, can be used to identify open reading frames (ORFs) within the genome that contain homology to ORFs from other organisms. Homology between sequences of interest can be determined using the local homology algorithm of Smith and Waterman, 1981, as well as the BestFit program (Rechid et al., 1989), and the FastDB algorithm (FastDB, 1988; described in *Current Methods in Sequence Comparison and Analysis, Macromolecule Sequencing and Synthesis, Selected Methods and Applications*, pp. 127-149, 1988, Alan R. Liss, Inc).

[0198] Alignment programs that permit gaps in the sequence include Clustalw (Thompson et al., 1994), FASTA3 (Pearson, 2000) Align0 (Myers and Miller, 1988), and Toffee (Notredame et al., 2000). Other methods for comparing and aligning nucleotide and protein sequences include, for example, BLASTX (NCBI), the Wise package (Birney and Durbin, 2000), and FASTX (Pearson, 2000). These algorithms determine sequence homology between nucleotide and protein sequences without translating the nucleotide sequences into protein sequences. Other techniques for alignment are also known in the art (Doolittle, et al., 1996; BLAST, available from the National Center for Biotechnology Information; FASTA, available in the Genetics Computing Group (GCG) package, from Madison, Wisconsin, USA, a wholly owned subsidiary of Oxford Molecular Group, Inc.; Schlessinger, 1988a; Schlessinger, 1988b; and Needleman and Wunch, 1970).

[0199] Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, etc. The reference sequence is usually at least about 18 nt long, at least about 30 nt long, or may extend to the complete sequence that is being compared.

[0200] One parameter for determining percent sequence identity is the percentage of the alignment in the region of strongest alignment between a target and a query sequence. Methods for determining this percentage involve, for example, counting the number of aligned bases of a query sequence in the region of strongest alignment and dividing this number by the total number of bases in the region. For example, 10 matches divided by 11 total residues gives a percent sequence identity of approximately 90.9%. The length of the aligned region is typically at least about 55%, at least about 58%, or at least about 60% of the total sequence length, and can be as great as about 62%, as great as about 64%, and even as great as about 66% of the total sequence length.

[0201] The present invention includes human and mouse polynucleotide and polypeptide sequences that are at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% homologous to the sequences in the appendices, based on using the method of determining sequence identity with the insertion of gaps to detect the maximum degree of sequence identity. In other embodiments of interest, homology will be at least about 80%, at least about 85%, or as high as about 90%.

[0202] A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. One format for an output means ranks the relative expression levels of different polynucleotides. Such presentation provides a skilled artisan with a ranking of relative expression levels to determine a gene expression profile.

[0203] As discussed above, the library of the invention also encompasses biochemical libraries of the polynucleotides shown in SEQ. ID. NOS. 1-187 and 375-484, *e.g.*, collections of nucleic acids representing the provided polynucleotides. The biochemical libraries can take a variety of forms, *e.g.*, a solution of cDNAs, a pattern of probe nucleic acids stably associated with a surface of a solid support (*i.e.*, an array) and the like. Of particular interest are nucleic acid arrays in which one or more of the polynucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484 is represented on the array. A variety of different array formats have been developed and are known to those of skill in the art. The arrays of the subject invention find use in a variety of applications, including gene expression analysis, drug screening, mutation analysis, and the like, as disclosed in the herein-listed exemplary patent documents.

[0204] In addition to the above nucleic acid libraries, analogous libraries of polypeptides are also provided, where the polypeptides of the library will represent at least a portion of the polypeptides encoded by a gene corresponding to one or more of the sequences shown in SEQ. ID. NOS. 1-187 and 375-484.

[0205] Further, analogous libraries of antibodies are also provided, where the libraries comprise antibodies or fragments thereof that specifically bind to at least a portion of at least one of the subject polypeptides. Further, antibody libraries may comprise antibodies or fragments thereof that specifically inhibit binding of a subject polypeptide to its ligand or substrate, or that specifically inhibit binding of a subject polypeptide as a substrate to another molecule. Moreover, corresponding nucleic acid libraries are also provided, comprising polynucleotide sequences that encode the antibodies or antibody fragments described above.

Polypeptides*Sequences*

[0206] This invention provides novel polypeptides, and related polypeptide compositions. The novel polypeptides of the invention encompass proteins with amino acid sequences as shown in SEQ. ID. NOS. 188-374, or encoded by the nucleic acids having nucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484. The subject polypeptides are human polypeptides, fragments thereof, variants (such as splice variants), homologs from other species, and derivatives thereof. In particular embodiments, a polypeptide of the invention has an amino acid sequence substantially identical to the sequence of any polypeptide encoded by a polynucleotide sequence shown in SEQ. ID. NOS. 1-187 and 375-484.

[0207] These polypeptides may reside within the cell, or extracellularly. They may be secreted from the cell, reside in the cytoplasm, in the membranes, or in any of the intracellular organelles, including the nucleus, mitochondria, ribosomes, or storage granules.

[0208] In some embodiments, the present novel polypeptide modulates the cells or tissues of animals, particularly humans, such as, for example, by stimulating, enhancing or inhibiting T or B cell function or the function of other hematopoietic cells or bone marrow cells; modulates adult or embryonic stem cell or precursor cell growth or differentiation; modulates cell function or activity of neuronal cells or other cells of the CNS, heart cells, liver cells, kidney cells, lung cells, pancreatic cells, gastrointestinal cells, spleen cells, breast cells, prostate cells, ovarian cells, and the like.

[0209] In some embodiments, a subject polypeptide is present as a multimer. Multimers include homodimers, homotrimers, homotetramers, and multimers that include more than four monomeric units. Multimers also include heteromultimers, *e.g.*, heterodimers, heterotrimers, heterotetramers, etc. where the subject polypeptide is present in a complex with proteins other than the subject polypeptide. Where the multimer is a heteromultimer, the subject polypeptide can be present in a 1:1 ratio, a 1:2 ratio, a 2:1 ratio, or other ratio, with the other protein(s).

[0210] In addition to the above specifically listed proteins, polypeptides from other species are also provided, including mammals, such as: primates, rodents, *e.g.*, mice, rats, hamsters, guinea pigs; domestic animals, *e.g.*, sheep, pig, horse, cow, goat, rabbit, dog, cat; and humans, as well as non-mammalian species, *e.g.*, avian, reptile and amphibian, insect, crustacean, fish, plant, fungus, and protozoa.

[0211] Also provided are polypeptides that are substantially identical to the at least one amino acid sequence shown in the appendices, or a fragment thereof, whereby substantially identical is meant that the protein has an amino acid sequence identity to the reference sequence of at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97%, at least about 98%, or at least about 99%.

[0212] The proteins of the subject invention (*e.g.*, polypeptides encoded by the nucleotide sequences shown in SEQ. ID. NOS. 1-187 and 375-484, and polypeptide sequences shown in SEQ. ID. NOS. 188-374) have been separated from their naturally occurring environment and are present in a non-naturally occurring environment. In certain embodiments, the proteins are present in a composition where they are more concentrated than in their naturally occurring environment. For example, purified polypeptides are provided.

[0213] In addition to naturally occurring proteins, polypeptides that vary from naturally occurring forms are also provided. Fusion proteins can comprise a subject polypeptide, or fragment thereof, and a polypeptide other than a subject polypeptide ("the fusion partner") fused in-frame at the N-terminus and/or C-terminus of the subject polypeptide, or internally to the subject polypeptide.

[0214] Suitable fusion partners include, but are not limited to, immunologically detectable proteins (*e.g.*, epitope tags, such as hemagglutinin, FLAG, and c-myc); polypeptides that provide a detectable signal or that serve as detectable markers (*e.g.*, a fluorescent protein, *e.g.*, a green fluorescent protein, a fluorescent protein from an Anthozoan species; β -galactosidase; luciferase; cre recombinase; and the like); polypeptides that provide a catalytic function or induce a cellular response; polypeptides that provide for secretion of the fusion protein from a eukaryotic cell; polypeptides that provide for secretion of the fusion protein from a prokaryotic cell; polypeptides that provide for binding to metal ions (*e.g.*, His_n, where n = 3-10, *e.g.*, 6His) and structural proteins. Fusion partners can also be those that are able to stabilize the present polypeptide, such as polyethylene glycol ("PEG") and a fragment of an immunoglobulin, such as the Fc fragment of IgG, IgE, IgA, IgM, and/or IgD.

[0215] Detection methods are chosen based on the detectable fusion partner. For example, where the fusion partner provides an immunologically recognizable epitope, an epitope-specific antibody can be used to quantitatively detect the level of polypeptide. In some embodiments, the fusion partner provides a detectable signal, and in these embodiments, the detection method is chosen based on the type of signal generated by the

fusion partner. For example, where the fusion partner is a fluorescent protein, fluorescence is measured.

[0216] Where the fusion partner is an enzyme that yields a detectable product, the product can be detected using an appropriate means. For example, β -galactosidase can, depending on the substrate, yield a colored product that can be detected with a spectrophotometer, and the fluorescent protein luciferase can yield a luminescent product detectable with a luminometer.

[0217] In some embodiments, a polypeptide of the invention comprises at least about 5, at least about 8, at least about 10, at least about 15, at least about 18, at least about 20, at least about 25, at least about 30, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 450, at least about 500, at least about 550, at least about 600, at least about 650, at least about 700, at least about 750, at least about 800, at least about 850, at least about 900, at least about 950, or at least about 1000 contiguous amino acid residues of at least one of the sequences according to SEQ. ID. NOS. 188-374, up to and including the entire amino acid sequence.

[0218] Fragments of the subject polypeptides, as well as polypeptides comprising such fragments, are also provided. Fragments of polypeptides of interest will typically be at least about 5, at least about 8, at least about 10, at least about 15, at least about 18, at least about 20, at least about 25, at least about 30, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, or at least 300 aa in length or longer, where the fragment will have a stretch of amino acids that is identical to the subject protein of at least about 5, at least about 8, at least about 10, at least about 15, at least about 18, at least about 20; at least about 25, at least about 30, or at least about 50 aa in length.

[0219] In some embodiments, fragments exhibit one or more activities associated with a corresponding naturally occurring polypeptide. Fragments find utility in generating antibodies to the full-length polypeptide; and in methods of screening for candidate agents that bind to and/or modulate polypeptide activity. Specific fragments of interest include those with enzymatic activity, those with biological activity including the ability to serve as an epitope or immunogen, and fragments that bind to other proteins or to nucleic acids.

[0220] The invention provides polypeptides comprising such fragments, including, *e.g.*, fusion polypeptides comprising a subject polypeptide fragment fused in frame (directly or indirectly) to another protein (the "fusion partner"), such as the signal peptide of one

protein being fused to the mature polypeptide of another protein. Such fusion proteins are typically made by linking the encoding polynucleotides together in a vector or cassette. Suitable fusion partners include, but are not limited to, immunologically detectable proteins (*e.g.*, epitope tags, such as hemagglutinin, FLAG, and c-myc); polypeptides that provide a detectable signal or that serve as detectable markers (*e.g.*, a fluorescent protein, *e.g.*, a green fluorescent protein, a fluorescent protein from an Anthozoan species; β -galactosidase; luciferase; cre recombinase); polypeptides that provide a catalytic function or induce a cellular response; polypeptides that provide for secretion of the fusion protein from a eukaryotic cell; polypeptides that provide for secretion of the fusion protein from a prokaryotic cell; polypeptides that provide for binding to metal ions (*e.g.*, His_n, where n = 3-10, *e.g.*, 6His) and structural proteins. Fusion partners can also be those that are able to stabilize the present polypeptide, such as polyethylene glycol ("PEG") and a fragment of an immunoglobulin, such as the Fc fragment of IgG, IgE, IgA, IgM, and/or IgD.

Polypeptide Preparation.

[0221] Polypeptides of the invention can be obtained from naturally-occurring sources or produced synthetically. The sources of naturally occurring polypeptides will generally depend on the species from which the protein is to be derived, *i.e.*, the proteins will be derived from biological sources that express the proteins. The subject proteins can also be derived from synthetic means, *e.g.*, by expressing a recombinant gene encoding a protein of interest in a suitable system or host or enhancing endogenous expression, as described in more detail above. Further, small peptides can be synthesized in the laboratory by techniques well known in the art.

[0222] In all cases, the product can be recovered by any appropriate means known in the art. For example, convenient protein purification procedures can be employed (*e.g.*, see Guide to Protein Purification, Deutscher et al., 1990). That is, a lysate can be prepared from the original source, (*e.g.*, a cell expressing endogenous polypeptide, or a cell comprising the expression vector expressing the polypeptide(s)), and purified using HPLC, exclusion chromatography, gel electrophoresis, or affinity chromatography, and the like.

[0223] The invention thus also provides methods of producing polypeptides. Briefly, the methods generally involve introducing a nucleic acid construct into a host cell *in vitro* and culturing the host cell under conditions suitable for expression, then harvesting the polypeptide, either from the culture medium or from the host cell, (*e.g.*, by disrupting the host cell), or both, as described in detail above. The invention also provides methods of producing

a polypeptide using cell-free *in vitro* transcription/translation methods, which are well known in the art, also as provided above.

[0224] Moreover, the invention provides polypeptides, including polypeptide fragments, as targets for therapeutic intervention, including use in screening assays, for identifying agents that modulate polypeptide level and/or activity, and as targets for antibody and small molecule therapeutics, for example, in the treatment of disorders.

Protein Families

[0225] The sequences of the invention encompass a variety of different types of nucleic acids and polypeptides with different structures and functions. They can encode or comprise polypeptides belonging to different protein families ("Pfam"). The "Pfam" system is an organization of protein sequence classification and analysis, based on conserved protein domains; it can be publicly accessed in a number of ways, for example, at <http://pfam.wustl.edu>. Protein domains are portions of proteins that have a tertiary structure and sometimes have enzymatic or binding activities; multiple domains can be connected by flexible polypeptide regions within a protein. Pfam domains can comprise the N-terminus or the C-terminus of a protein, or can be situated at any point in between. The Pfam system identifies protein families based on these domains and provides an annotated, searchable database that classifies proteins into families (Bateman et al., 2002).

[0226] Sequences of the invention can encode or be comprised of more than one Pfam. Sequences encompassed by the invention include, but are not limited to, the polypeptide and polynucleotide sequences of the molecules shown in the appendices and corresponding molecular sequences found at all developmental stages of an organism. Sequences of the invention can comprise genes or gene segments designated in the appendices, and their gene products, *i.e.*, RNA and polypeptides. They also include variants of those presented in the appendices that are present in the normal physiological state, *e.g.*, variant alleles such as SNPs, splice variants, as well as variants that are affected in pathological states, such as disease-related mutations or sequences with alterations that lead to pathology, and variants with conservative amino acid changes. Some sequences of the invention are categorized below with respect to one or more protein family. Any given sequence can belong to one or more than one category.

[0227] Some sequences of the invention can encode or be comprised of reverse transcriptase (rvt) domains, which are involved in RNA-directed DNA polymerase activity, an enzymatic activity that uses an RNA template to produce DNA for integration into a host

genome (<http://pfam.wustl.edu/cgi-bin/getdesc?name=rvt>). Some sequences of the invention can encode or be comprised of transposase 1 (Transposase_1) domains, which are characterized by sequences that can excise and/or insert mobile genetic elements such as transposons or insertion sequences, for example, the mariner protein (http://pfam.wustl.edu/cgi-bin/getdesc?name=Transposase_1). Some sequences of the invention can encode or be comprised of gag gene protein p24 (core nucleocapsid protein or gag_p24) domains, which comprise the proteins of the nucleocapsid shell around the RNA of a retrovirus. The gag_p24, or p24, protein from retroviruses forms the inner protein layer of the nucleocapsid, and performs orchestrated tasks during the assembly, budding, maturation, and infection stages of the viral replication cycle. Gag precursors also function during viral assembly to selectively bind and package strands of genomic RNA. ELISA tests for p24 are among the most commonly used methods to demonstrate virus replication both *in vivo* and *in vitro*. (http://pfam.wustl.edu/cgi-bin/getdesc?name=Gag_p24).

[0228] Some sequences of the invention can encode or be comprised of integrase core domains, which mediate the integration of a DNA copy of a viral genome into a host chromosome; *e.g.*, HIV integrase catalyses the incorporation of virally derived DNA into the human genome, presenting a target for the development of new therapeutics for the treatment of AIDS (<http://pfam.wustl.edu/cgi-bin/getdesc?name=rve>). Some sequences of the invention can encode or be comprised of RNase H domains, which are nucleases specific for the RNA strand of an RNA-DNA hybrid that cleaves phosphodiester bonds to produce molecules with 3'-OH and 5'-PO₄ ends (<http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00075>). Antisense deoxyoligonucleotides that trigger RNase H activity can thus be used as cancer therapeutic agents (Crooke, 1996; Curcio et al., 1997).

Methods

[0229] The present invention provides methods of producing a subject polypeptide and provides antibodies that specifically bind to a subject polypeptide. The present invention further provides screening methods for identifying agents that modulate a level or an activity of a subject polypeptide or polynucleotide. The present invention thus also provides agents that modulate a level or an activity of a subject polypeptide or polynucleotide, as well as compositions, including pharmaceutical compositions, comprising a subject agent.

[0230] The present invention further provides methods for treating disorders such as, for example, cancer and other proliferative disorders, inflammatory and immune disorders, metabolic disorders, and bacterial or viral disorders.

Diagnostic and Therapeutic Applications

Screening and Diagnostic Methods

1. Identifying Biological Molecules that Interact with a Polypeptide

[0231] Formation of a binding complex between a subject polypeptide and an interacting polypeptide or other macromolecule (*e.g.*, DNA, RNA, lipids, polysaccharides, and the like) can be detected using any known method. Suitable methods include: a yeast two-hybrid system (Zhu et al., 1997; Fields and Song, 1989; U.S. Pat. No. 5,283,173; Chien et al. 1991); a mammalian cell two-hybrid method; a fluorescence resonance energy transfer (FRET) assay; a bioluminescence resonance energy transfer (BRET) assay; a fluorescence quenching assay; a fluorescence anisotropy assay (Jameson and Sawyer, 1995); an immunological assay; and an assay involving binding of a detectably labeled protein to an immobilized protein.

[0232] Immunological assays, and assays involving binding of a detectably labeled protein to an immobilized protein can be performed in a variety of ways. For example, immunoprecipitation assays can be designed such that the complex of protein and an interacting polypeptide is detected by precipitation with an antibody specific for either the protein or the interacting polypeptide.

[0233] FRET detects formation of a binding complex between a subject polypeptide and an interacting polypeptide. It involves the transfer of energy from a donor fluorophore in an excited state to a nearby acceptor fluorophore. For this transfer to take place, the donor and acceptor molecules must be in close proximity (*e.g.*, less than 10 nanometers apart, usually between 10 and 100 Å apart), and the emission spectra of the donor fluorophore must overlap the excitation spectra of the acceptor fluorophore. In these embodiments, a fluorescently labeled subject protein serves as a donor and/or acceptor in combination with a second fluorescent protein or dye.

[0234] Fluorescent proteins can be produced by generating a construct comprising a protein and a fluorescent fusion partner. These are well-known in the art, as described above, including green fluorescent protein (GFP), *i.e.*, a "humanized" version of a GFP, *e.g.*, wherein codons of the naturally-occurring nucleotide sequence are changed to more closely match human codon bias; a GFP derived from *Aequoria victoria* or a derivative thereof, *e.g.*, a "humanized" derivative such as Enhanced GFP, which are available commercially, *e.g.*, from Clontech, Inc.; other fluorescent mutants of a GFP from *Aequoria victoria*, *e.g.*, as described in U.S. Patent No. 6,066,476; 6,020,192; 5,985,577; 5,976,796; 5,968,750; 5,968,738;

5,958,713; 5,919,445; 5,874,304; a GFP from another species such as *Renilla reniformis*, *Renilla mulleri*, or *Ptilosarcus guernyi*, as previously described (WO 99/49019; Peelle et al., 2001), "humanized" recombinant GFP (hrGFP) (Stratagene[®]); any of a variety of fluorescent and colored proteins from Anthozoan species, (*e.g.*, Matz et al., 1999); as well as proteins labeled with other fluorescent dyes, fluorescein and its derivatives, *e.g.*, fluorescein isothiocyanate (FITC), 6-carboxyfluorescein (6-FAM), 6-carboxy-2',4',7',4',7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE); rhodamine dyes, *e.g.*, Texas red, phycoerythrin, tetramethylrhodamine, rhodamine, 6-carboxy-X-rhodamine (ROX); coumarin and its derivatives, *e.g.*, 7-amino-4-methylcoumarin, aminocoumarin; bodipy dyes, such as Bodipy FL; cascade blue; Oregon green; eosins and erythrosins; cyanine dyes, *e.g.*, allophycocyanin, Cy3, Cy5, and N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA); macrocyclic chelates of lanthanide ions, *e.g.*, quantum dye, etc; and chemiluminescent molecules, *e.g.*, luciferases.

[0235] Fluorescent subject proteins can also be generated by producing the subject protein in an auxotrophic strain of bacteria which requires addition of one or more amino acids in the medium for growth. A subject protein-encoding construct that provides for expression in bacterial cells is introduced into the auxotrophic strain, and the bacteria are cultured in the presence of a fluorescent amino acid, which is incorporated into the subject protein produced by the bacterium. The subject protein is then purified from the bacterial culture using standard methods for protein purification.

[0236] BRET is a protein-protein interaction assay based on energy transfer from a bioluminescent donor to a fluorescent acceptor protein. The BRET signal is measured by the ratio of the amount of light emitted by the acceptor to the amount of light emitted by the donor. The ratio of these two values increases as the two proteins are brought into proximity. The BRET assay has been described in the literature (U.S. Patent Nos. 6,020,192; 5,968,750; 5,874,304; Xu, et al. 1999). BRET assays can be performed by analyzing transfer between a bioluminescent donor protein and a fluorescent acceptor protein. Interaction between the donor and acceptor proteins can be monitored by a change in the ratio of light emitted by the bioluminescent and fluorescent proteins. In this application, the subject protein serves as donor and/or acceptor protein.

[0237] Fluorescence anisotropy is a measurement of the rotational mobility of a multi-molecular complex. It can be used to generate information about the binding of one

molecule to another, including the affinity and specificity of binding sites. It can be applied to polypeptides or nucleic acids of the present invention.

[0238] Fluorescence quenching measurements are useful in detecting protein multimerization, such as where the subject protein interacts with at least a second protein and, for example, where multimerization interaction is affected by a test agent. As used herein, the term "multimerization" refers to formation of dimers, trimers, tetramers, and higher multimers of the subject protein. Whether a subject protein forms a complex with one or more additional protein molecules can be determined using any known assay, including assays as described above for interacting proteins. Formation of multimers can also be detected using non-denaturing gel electrophoresis, where multimerized subject protein migrates more slowly than monomeric subject protein. Formation of multimers can also be detected using fluorescence quenching techniques.

[0239] Formation of multimers can also be detected by analytical ultracentrifugation, for example through glycerol or sucrose gradients, and subsequent visualization of a subject protein in gradient fractions by Western blotting or staining of SDS-polyacrylamide gels. Multimers are expected to sediment at defined positions in such gradients. Formation of multimers can also be detected using analytical gel filtration, *e.g.*, in HPLC or FPLC systems, *e.g.*, on columns such as Superdex 200 (Pharmacia Amersham Inc.). Multimers run at defined positions on these columns, and fractions can be analyzed as above. The columns are highly reproducible, allowing one to relate the number and position of peaks directly to the multimerization status of the protein.

2. Detecting mRNA Levels and Monitoring Gene Expression

[0240] The present invention provides methods for detecting the presence of mRNA in a biological sample. The methods can be used, for example, to assess whether a test compound affects gene expression, either directly or indirectly. The present invention provides diagnostic methods to compare the abundance of a nucleic acid with that of a control value, either qualitatively or quantitatively, and to relate the value to a normal or abnormal expression pattern.

[0241] Methods of measuring mRNA levels are known in the art (Pietu, 1996; Zhao, 1995; Soares, 1997; Raval, 1994; Chalifour, 1994; Stolz, 1996; Hong, 1982; McGraw, 1984; WO 97/27317). These methods generally comprise contacting a sample with a polynucleotide of the invention under conditions that allow hybridization and detecting hybridization, if any, as an indication of the presence of the polynucleotide of interest.

Appropriate controls include the use of a sample lacking the polynucleotide mRNA of interest, or the use of a labeled polynucleotide of the same "sense" as a polynucleotide mRNA of interest. Detection can be accomplished by any known method, including, but not limited to, *in situ* hybridization, PCR, RT-PCR, and "Northern" or RNA blotting, or combinations of such techniques, using a suitably labeled subject polynucleotide. A variety of labels and labeling methods for polynucleotides are known in the art and can be used in the assay methods of the invention. A common method employed is use of microarrays which can be purchased or customized, for example, through conventional vendors such as Affymetrix.

[0242] In some embodiments, the methods involve generating a cDNA copy of an mRNA molecule in a biological sample, and amplifying the cDNA using an isolated primer pairs as described above, *i.e.*, a set of two nucleic acid molecules that serve as forward and reverse primers in an amplification reaction (*e.g.*, a polymerase chain reaction). The primer pairs are chosen to specifically amplify a cDNA copy of an mRNA encoding a polypeptide. A detectable label can be included in the amplification reaction, as provided above. Methods using PCR amplification can be performed on the DNA from a single cell, although it is convenient to use at least about 10^5 cells.

[0243] The present invention provides methods for monitoring gene expression. Changes in a promoter or enhancer sequence that can affect gene expression can be examined in light of expression levels of the normal allele by various methods known in the art. Methods for determining promoter or enhancer strength include quantifying the expressed natural protein, and inserting the variant control element into a vector with a quantitative reporter gene such as β -galactosidase, luciferase, or chloramphenicol acetyltransferase (CAT).

3. *Detecting Polymorphisms and Mutations*

[0244] Biochemical studies can determine whether a sequence polymorphism in a coding region or control region is associated with disease. Disease-associated polymorphisms can include deletion or truncation of the gene, mutations that alter expression level, or mutations that affect protein function, etc. A number of methods are available to analyze nucleic acids for the presence of a specific sequence, *e.g.*, a disease associated polymorphism. Genomic DNA can be used when large amounts of DNA are available. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis. Cells that express the gene provide a source of mRNA, which can be

assayed directly or reverse transcribed into cDNA for analysis. The nucleic acid can be amplified by conventional techniques, *i.e.*, PCR, to provide sufficient amounts for analysis. (Saiki et al., 1988; Sambrook et al., 1989, pp.14.2-14.33). Alternatively, various methods are known in the art that utilize oligonucleotide ligation as a means of detecting polymorphisms (Riley et al., 1990; Delahunty et al., 1996).

[0245] The sample nucleic acid, *e.g.*, an amplified or cloned fragment, is analyzed by one of a number of methods known in the art. The nucleic acid can be sequenced by dideoxy nucleotide sequencing, or other methods, and the sequence of bases compared to a wild-type sequence. Hybridization with the variant sequence can also be used to determine its presence, *e.g.*, by Southern blots, dot blots, etc. The hybridization pattern of a control and variant sequence to an array of oligonucleotide probes immobilized on a solid support, as described in US Pat. No. 5,445,934, or WO 95/35505, can also be used as a means of detecting the presence of variant sequences. Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices can detect variation as alterations in electrophoretic mobility resulting from conformational changes created by DNA sequence alterations. Alternatively, where a polymorphism creates or destroys a recognition site for a restriction endonuclease, the sample can be digested with that endonuclease, and the products fractionated according to their size to determine whether the fragment was digested. Fractionation can be performed by gel or capillary electrophoresis, for example with acrylamide or agarose gels.

[0246] Screening for mutations in a gene can be based on the functional or antigenic characteristics of the protein. Protein truncation assays are useful in detecting deletions that might affect the biological activity of the protein. Various immunoassays designed to detect polymorphisms in proteins can be used in screening. Where many diverse genetic mutations lead to a particular disease phenotype, functional protein assays have proven to be effective screening tools. The activity of the encoded protein can be determined by comparison with the wild-type protein.

4. Detecting and Monitoring Polypeptide Presence and Biological Activity

[0247] The present invention provides methods for detecting the presence and/or biological activity of a subject polypeptide in a biological sample. The assay used will be appropriate to the biological activity of the particular polypeptide. Thus, *e.g.*, where the biological activity is an enzymatic activity, the method will involve contacting the sample with an appropriate substrate, and detecting the product of the enzymatic reaction on the

substrate. Where the biological activity is binding to a second macromolecule, the assay detects protein-protein binding, protein-DNA binding, protein-carbohydrate binding, or protein-lipid binding, as appropriate, using well known assays. Where the biological activity is signal transduction (*e.g.*, transmission of a signal from outside the cell to inside the cell) or transport, an appropriate assay is used, such as measurement of intracellular calcium ion concentration, measurement of membrane conductance changes, or measurement of intracellular potassium ion concentration.

[0248] The present invention also provides methods for detecting the presence or measuring the level of a normal or abnormal polypeptide in a biological sample using a specific antibody. The methods generally comprise contacting the sample with a specific antibody and detecting binding between the antibody and molecules of the sample. Specific antibody binding, when compared to a suitable control, is an indication that a polypeptide of interest is present in the sample. Suitable controls include a sample known not to contain the polypeptide, and a sample contacted with a non-specific antibody, *e.g.*, an anti-idiotypic antibody.

[0249] A variety of methods to detect specific antibody-antigen interactions are known in the art, *e.g.*, standard immunohistological methods, immunoprecipitation, enzyme immunoassay, and radioimmunoassay. The specific antibody can be detectably labeled, either directly or indirectly, as described at length herein, and cells are permeabilized to stain cytoplasmic molecules. Briefly, antibodies are added to a cell sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody may be labeled with radioisotopes, enzymes, fluorescers, chemiluminescers, or other labels for direct detection.. Alternatively, specific-binding pairs may be used, involving, *e.g.*, a second stage antibody or reagent that is detectably-labeled, as described above. Such reagents and their methods of use are well known in the art

[0250] Alternatively, a biological sample can be brought into contact with an immobilized antibody on a solid support or carrier, such as nitrocellulose, that is capable of immobilizing cells, cell particles, or soluble proteins. The antibody can be attached (coupled) to an insoluble support, such as a polystyrene plate or a bead. After contacting the sample, the support can then be washed with suitable buffers, followed by contacting with a detectably-labeled specific antibody. Detection methods are known in the art and will be chosen as appropriate to the signal emitted by the detectable label. Detection is generally accomplished in comparison to suitable controls, and to appropriate standards.

[0251] The present invention further provides methods for detecting the presence and/or levels of enzymatic activity of a subject polypeptide in a biological sample. The methods generally involve contacting the sample with a substrate that yields a detectable product upon being acted upon by a subject polypeptide, and detecting a product of the enzymatic reaction. Further, polypeptides that are subsets of the complete sequences of the subject proteins may be used to identify and investigate parts of the protein important for function.

[0252] The present invention further includes methods for monitoring activity of a polypeptide through observation of phenotypic changes in a cell containing such polypeptide, such as growth or differentiation, or the ability of such a cell to secrete a molecule that can be detected, such as through chemical methods or through its effect on another cell, such as cell activation.

5. Modulating mRNA and Peptides in Biological Samples

[0253] The present invention provides screening methods for identifying agents that modulate the level of a mRNA molecule of the invention, agents that modulate the level of a polypeptide of the invention, and agents that modulate the biological activity of a polypeptide of the invention. In some embodiments, the assay is cell-free; in others, it is cell-based. Where the screening assay is a binding assay, one or more of the molecules can be joined to a label, where the label can directly or indirectly provide a detectable signal.

[0254] As discussed above, the invention encompasses endogenous polynucleotides of the invention that encode mRNA and/or polypeptides of interest. Again as discussed previously, the invention also encompasses exogenous polynucleotides that encode mRNA or polypeptides of the invention. For example, the polynucleotide can reside within a recombinant vector which is introduced into the cell. For example, a recombinant vector can comprise an isolated transcriptional regulatory sequence which is associated in nature with a nucleic acid, such as a promoter sequence operably linked to sequences coding for a polypeptide of the invention; or the transcriptional control sequences can be operably linked to coding sequences for a polypeptide fusion protein comprising a polypeptide of the invention fused to a polypeptide that facilitates detection.

[0255] In these embodiments, the candidate agent is combined with a cell possessing a polynucleotide transcriptional regulatory element operably linked to a polypeptide-coding sequence of interest, *e.g.*, a subject cDNA or its genomic component; and determining the

agent's effect on polynucleotide expression, as measured, for example by the level of mRNA, polypeptide, or fusion polypeptide.

[0256] In other embodiments, for example, a recombinant vector can comprise an isolated polynucleotide transcriptional regulatory sequence, such as a promoter sequence, operably linked to a reporter gene (*e.g.*, β -galactosidase, CAT, luciferase, or other gene that can be easily assayed for expression). In these embodiments, the method for identifying an agent that modulates a level of expression of a polynucleotide in a cell comprises combining a candidate agent with a cell comprising a transcriptional regulatory element operably linked to a reporter gene; and determining the effect of said agent on reporter gene expression.

[0257] Known methods of measuring mRNA levels can be used to identify agents that modulate mRNA levels, including, but not limited to, PCR with detectably-labeled primers. Similarly, agents that modulate polypeptide levels can be identified using standard methods for determining polypeptide levels, including, but not limited to an immunoassay such as ELISA with detectably-labeled antibodies.

[0258] A wide variety of cell-based assays can also be used to identify agents that modulate eukaryotic or prokaryotic mRNA and/or polypeptide levels. Examples include transformed cells that over-express a cDNA construct and cells transformed with a polynucleotide of interest associated with an endogenously-associated promoter operably linked to a reporter gene. A control sample would comprise, for example, the same cell lacking the candidate agent. Expression levels are measured and compared in the test and control samples.

[0259] The cells used in the assay are usually mammalian cells, including, but not limited to, rodent cells and human cells. The cells can be primary cell cultures or can be immortalized cell lines. Cell-based assays generally comprise the steps of contacting the cell with a test agent, forming a test sample, and, after a suitable time, assessing the agent's effect on macromolecule expression. That is, the mammalian cell line is transformed or transfected with a construct that results in expression of the polynucleotide, the cell is contacted with a test agent, and then mRNA or polypeptide levels are detected and measured using conventional assays.

[0260] A suitable period of time for contacting the agent with the cell can be determined empirically, and is generally a time sufficient to allow entry of the agent into the cell and to allow the agent to have a measurable effect on subject mRNA and/or polypeptide levels. Generally, a suitable time is between about 10 minutes and about 24 hours, including

about 1 to about 8 hours. Alternatively, incubation periods may be between about 0.1 and about 1 hour, selected for example for optimum activity or to facilitate rapid high-throughput screening. Where the polypeptide is expressed on the cell surface, however, a shorter length of time may be sufficient. Incubations are performed at any suitable temperature, *i.e.*, between about 4°C and about 40°C. The contact and incubation steps can be followed by a washing step to remove unbound components, *i.e.*, a label that would give rise to a background signal during subsequent detection of specifically-bound complexes.

[0261] A variety of assay configurations and protocols are known in the art. For example, one of the components can be bound to a solid support, and the remaining components contacted with the support bound component. Remaining components may be added at different times or at substantially the same time. Further, where the interacting protein is a second subject protein, the effect of the test agent on binding can be determined by determining the effect on multimerization of the subject protein.

[0262] The present invention further provides methods of identifying agents that modulate a biological activity of a polypeptide of the invention. The method generally comprises contacting a test agent with a sample containing a subject polypeptide and assaying a biological activity of the subject polypeptide in the presence of the test agent. An increase or a decrease in the assayed biological activity in comparison to the activity in a suitable control (*e.g.*, a sample comprising a subject polypeptide in the absence of the test agent) is an indication that the substance modulates a biological activity of the subject polypeptide. The mixture of components is added in any order that provides for the requisite interaction..

[0263] External and internal processes that can affect modulation of a macromolecule of the invention include, but are not limited to, infection of a cell by a microorganism, including, but not limited to, a bacterium (*e.g.*, *Mycobacterium* spp., *Shigella*, or *Chlamydia*), a protozoan (*e.g.*, *Trypanosoma* spp., *Plasmodium* spp., or *Toxoplasma* spp.), a fungus, a yeast (*e.g.*, *Candida* spp.), or a virus (including viruses that infect mammalian cells, such as human immunodeficiency virus, foot and mouth disease virus, Epstein-Barr virus, and viruses that infect plant cells); change in pH of the medium in which a cell is maintained or a change in internal pH; excessive heat relative to the normal range for the cell or the multicellular organism; excessive cold relative to the normal range for the cell or the multicellular organism; an effector molecule such as a hormone, a cytokine, a chemokine, a neurotransmitter; an ingested or applied drug; a ligand for a cell-surface receptor; a ligand for

a receptor that exists internally in a cell, *e.g.*, a nuclear receptor; hypoxia; light; dark; sleep patterns; electrical charge; ion concentration of the medium in which a cell is maintained or an internal ion concentration, exemplary ions including sodium ions, potassium ions, chloride ions, calcium ions, and the like; presence or absence of a nutrient; metal ions; a transcription factor; mitogens, including, but not limited to, lipopolysaccharide (LPS), pokeweed mitogen; antigens; a tumor suppressor; and cell-cell contact and must be taken into consideration in the screening assay.

[0264] A variety of other reagents can be included in the screening assay. These include salts, neutral proteins, *e.g.*, albumin, detergents, and other compounds that facilitate optimal binding and/or reduce non-specific or background interactions. Reagents that improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, or anti-microbial agents, etc., can be used.

[0265] Accordingly, the present invention provides a method for identifying an agent, particularly a biologically active agent that modulates the level of expression of a nucleic acid in a cell, the method comprising: combining a candidate agent to be tested with a cell comprising a nucleic acid that encodes a polypeptide, and determining the agent's effect on polypeptide expression.

[0266] Some embodiments will detect agents that decrease the biological activity of a molecule of the invention. Maximal inhibition of the activity is not always necessary, or even desired, in every instance to achieve a therapeutic effect. Agents that decrease a biological activity can find use in treating disorders associated with the biological activity of the molecule. Alternatively, some embodiments will detect agents that increase a biological activity. Agents that increase a biological activity of a molecule of the invention can find use in treating disorders associated with a deficiency in the biological activity. Agents that increase or decrease a biological activity of a molecule of the invention can be selected for further study, and assessed for physiological attributes, *i.e.*, cellular availability, cytotoxicity, or biocompatibility, and optimized as required. For example, a candidate agent is assessed for any cytotoxic activity it may exhibit toward the cell used in the assay using well-known assays, such as trypan blue dye exclusion, an MTT ([3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2 H-tetrazolium bromide]) assay, and the like.

[0267] A variety of different candidate agents can be screened by the above methods. Candidate agents encompass numerous chemical classes, as described above.

[0268] Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. Numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and oligopeptides. For example, random peptide libraries obtained by yeast two-hybrid screens (Xu et al., 1997), phage libraries (Hoogenboom et al., 1998), or chemically generated libraries. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced, including antibodies produced upon immunization of an animal with subject polypeptides, or fragments thereof, or with the encoding polynucleotides. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and can be used to produce combinatorial libraries. Further, known pharmacological agents can be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, and amidification, etc, to produce structural analogs.

6. Kits

[0269] The present invention provides methods for diagnosing disease states based on the detected presence and/or level of polynucleotide or polypeptide in a biological sample, and/or the detected presence and/or level of biological activity of the polynucleotide or polypeptide. These detection methods can be provided as part of a kit. Thus, the invention further provides kits for detecting the presence and/or a level of a polynucleotide or polypeptide in a biological sample and/or or the detected presence and/or level of biological activity of the polynucleotide or polypeptide. Procedures using these kits can be performed by clinical laboratories, experimental laboratories, medical practitioners, or private individuals.

[0270] The kits of the invention will comprise a molecule of the invention. The kits for detecting a polynucleotide will also comprise a moiety that specifically hybridizes to a polynucleotide of the invention. The polynucleotide molecule can be of any length. For example, it can comprise a polynucleotide of at least 6, at least 7, at least 8, or at least 9 contiguous nucleotides of a molecule of the invention. Kits of the invention for detecting a subject polypeptide will comprise a moiety that specifically binds to a polypeptide of the invention; the moiety includes, but is not limited to, a polypeptide-specific antibody.

[0271] The kits are useful in diagnostic applications. For example, the kit is useful to determine whether a given DNA sample isolated from an individual comprises an expressed nucleic acid, a polymorphism, or other variant.

[0272] Kits for detecting polynucleotides comprise a pair of nucleic acids in a suitable storage medium, *e.g.*, a buffered solution, in a suitable container. The pair of isolated nucleic acid molecules serve as primers in an amplification reaction (*e.g.*, a polymerase chain reaction). The kit can further include additional buffers, reagents for polymerase chain reaction (*e.g.*, deoxynucleotide triphosphates (dNTP), a thermostable DNA polymerase, a solution containing Mg^{2+} ions (*e.g.*, $MgCl_2$), and other components well known to those skilled in the art for carrying out a polymerase chain reaction). The kit can further include instructions for use, which may be provided in a variety of forms, *e.g.*, printed information, or compact disc, and the like. The kit may further include reagents necessary to extract DNA from a biological sample and reagents for generating a cDNA copy of an mRNA. The kit may optionally provide additional useful components, including, but not limited to, buffers, developing reagents, labels, reacting surfaces, means for detections, control samples, standards, and interpretive information.

[0273] In some embodiments, a kit of the invention for detecting a polynucleotide, such as an mRNA encoding a polypeptide, comprises a pair of nucleic acids that function as "forward" and "reverse" primers that specifically amplify a cDNA copy of the mRNA. The "forward" and "reverse" primers are provided as a pair of isolated nucleic acid molecules, each from about 10 to about 200 nucleotides in length, the first nucleic acid molecule of the pair comprising a sequence of at least about 10 contiguous nucleotides having 100% sequence identity to a nucleic acid sequence shown in SEQ. ID. NOS. 1-187 and 375-484, and the second nucleic acid molecule of the pair comprising a sequence of at least about 10 contiguous nucleotides having 100% sequence identity to the reverse complement of a nucleic acid sequence shown in SEQ. ID. NOS. 1-187 and 375-484, wherein the sequence of the second nucleic acid molecule is located 3' of the nucleic acid sequence of the first nucleic acid molecule. The primer nucleic acids are prepared using any known method, *e.g.*, automated synthesis. In some embodiments, one or both members of the pair of nucleic acid molecules comprise a detectable label.

[0274] Where the kit provides for polypeptide detection, it can include one or more specific antibodies. In some embodiments, the antibody specific to the polypeptide is detectably labeled. In other embodiments, the antibody specific to the polypeptide is not

labeled; instead, a second, detectably-labeled antibody is provided that binds to the specific antibody. The kit may further include blocking reagents, buffers, and reagents for developing and/or detecting the detectable marker. The kit may further include instructions for use, controls, and interpretive information.

[0275] Where the kit provides for detecting enzymatic activity, it includes a substrate that provides for a detectable product when acted upon by a polypeptide of interest. The kit may further include reagents necessary to detect and develop the detectable marker.

[0276] The present invention provides for kits with unit doses of an active agent. These agents are described in more detail below. In some embodiments, the agent is provided in oral or injectable doses. Such kits will comprise containers containing the unit doses and an informational package insert describing the use and attendant benefits of the drugs in treating a condition of interest.

Stem Cells

[0277] Stem cells are pluripotent or multipotent cells that generate maturing cells in multiple differentiation lineages. Embryonic stem cells are pluripotent; they can differentiate into any of the cells present in the adult. Multipotent cells have the ability to differentiate into more than one cell type. Organ-specific stem cells are multipotent; they can differentiate into any of the cells of the organ they inhabit.

[0278] When they divide *in vivo*, both pluripotent and multipotent stem cells can maintain their pluripotency or multipotency while giving rise to differentiated progeny. Thus, stem cells can produce replicas of themselves which are pluri- or multipotent, and are also able to differentiate into lineage-restricted committed progenitor cells. For example, hematopoietic stem cells, which are multipotent cells specifically able to form blood cells, can divide to produce replicate hematopoietic stem cells. They can also divide to produce more highly differentiated cells, which are precursors of blood cells. The precursors differentiate, sometimes through several generations of cells, into blood cells. A hematopoietic stem cell can also divide into a cell with the capacity to form, for example, a relatively undifferentiated cell that is committed to differentiate into, *i.e.*, granulocytes, erythrocytes, or another type of blood cell.

[0279] Stem cells can also reproduce and differentiate *in vitro*. Embryonic stem cells have been directed to differentiate into cardiac muscle cells *in vitro* and, alternatively, into early progenitors of neural stem cells, and then into mature neurons and glial cells *in vitro* (Trounson, 2002).

[0280] The study of stem cells has led to therapy for treating cancer in humans (Slavin et al., 2001). Stem cell therapy offers several advantages over traditional cancer therapies (Weissman, 2000). One advantage of stem cell therapy exists when used in conjunction with radiation therapy. In radiation therapy for cancer, the dose of radiation necessary to kill the cancer cells in an organ can also be sufficient to destroy the healthy cells of the organ. In combined stem cell and radiation therapy, an organ is first treated with sufficient radiation to destroy all of the cancer cells and most or all of the healthy cells, but then stem cells are infused to repopulate the organ. In the ensuing weeks, as the cancer cells and healthy cells die, the stem cells replace the healthy cells. Another advantage of this approach, compared to heterologous organ transplants, is that there is no risk of rejection, since stem cells do not provoke an immune response. A further advantage is that stem cells are inherently programmed to regulate their numbers and differentiation status, *i.e.*, once provided to the patient, the necessary number will differentiate, and the rest will remain undifferentiated (Weissman, 2000).

[0281] Stem cell therapy is also effective in treating autoimmune disease in humans. For example, immunosuppression in conjunction with stem-cell transplantation has induced remission in patients with refractory, severe rheumatic autoimmune disease (Van Laar and Tyndall, 2003). Patients with rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, and juvenile idiopathic arthritis have benefited from stem cell transplants (Van Laar and Tyndall, 2003).

[0282] Preclinical studies also suggest the potential of stem cell transplantation for the treatment of neural and muscular injuries and diseases, including those of the central nervous system, peripheral nervous system, and skeletal, cardiac and smooth muscle (Deasy and Huard, 2002). Stem cells transplanted into the bone marrow of mice migrate to the site of injured muscle and differentiate into new muscle cells. Thus, patients with myasthenia gravis, muscular dystrophies, amyotrophic lateral sclerosis, congestive heart failure, Parkinson's disease, and Alzheimer's disease may benefit from stem cell therapy (Henningson, 2003).

[0283] In addition to therapeutic uses, research using stem cells can provide useful information about normal stem cell function and the pathogenesis of disease. Stem cells derived from a patient with a genetic disease can provide a tool for studying that disease. To derive these stem cells, a somatic cell, *i.e.*, a cell that is not in the oocyte or spermatocyte lineage, is donated by the patient, and the nucleus is removed and transferred to an

unfertilized human oocyte. This nuclear transplant procedure produces, at the blastocyst stage of development, embryonic stem cells with the same set of genes as the patient with the genetic disease. Studying these cells, and their progeny *in vitro*, permits analysis of a specific model of the disease. For example, placing stem cells derived from a patient with a genetic disease under the control of various stem cell regulatory factors can elicit abnormal responses from the affected stem cells compared to stem cells derived from a healthy individual's somatic nucleus.

[0284] A specific type of stem cell, mesenchymal stem cells (MSC) are adult pluripotent stem cells present *in vivo* in the bone marrow, and able to travel to and populate different sites in the body. MSCs derive from a common progenitor cell, the multipotent adult progenitor cell (MAPC), that gives rise to other lineages as different as endothelium, endoderm, and ectoderm. MSCs self-renew by proliferation while maintaining their stem cell phenotype and give rise to the differentiated stromal cells which belong to the osteogenic, chondrogenic, adipogenic, myogenic and fibroblastic lineages. They are capable of differentiating into a number of different cell types, including osteocytes, chondrocytes, adipocytes, myocytes, and fibroblasts (Van Damme et al., 2002).

[0285] MSCs can be identified by their expression of thy-1 (CD90), endoglin (CD105), vascular cell adhesion molecule 1 (VCAM-1) (CD106), and hyaluronate receptor (CD44). MSC potential can be determined by the colony-forming unit fibroblast assay (CFU-F) (Jorgensen et al., 2003). They can be expanded in culture without losing their phenotype or multilineage potential (Jorgensen et al., 2003).

[0286] MSCs are present in tissues involved in autoimmune diseases. Since MSCs can target the affected organs in autoimmune diseases, they can be used as a vehicle to express therapeutic proteins (Jorgensen et al., 2003).

[0287] MSCs have been demonstrated to be present in mesenchymal-derived tissue, such as injured synovium in the presence of chronic inflammation (Jorgensen et al., 2003).

[0288] MSCs themselves display immunological properties; these properties can enhance their therapeutic potential (Jorgensen et al., 2003). For example, they suppress T-cell responses (Jorgensen et al., 2003). Systemically infused autologous MSCs delayed donor cell rejection from a control value of 7 days to a treated value of 11 days in a murine skin graft model (Jorgensen et al., 2003).

[0289] Thus, MSCs can be used in different therapeutic strategies, either as immunosuppressive agents, or as genetically engineered vehicles for therapeutics. Since

MSCs are not rejected, and they travel to injured mesenchymal tissues, they can be used as a means of transportation to deliver therapeutic proteins to a targeted area. The capacity of MSCs to adhere to matrix components favors their preferential homing to bone, lung, and cartilage when injected intravenously (Jorgensen et al., 2003). They also home preferentially to the bone marrow. Cell mediated therapy offers the advantage that the autologous MCS can be used in treatment.

[0290] MSCs have been genetically modified to express anti-inflammatory cytokines. They have been demonstrated to secrete several therapeutic heterologous proteins, such as human erythropoietin and factor IX (Jorgensen et al., 2003). Rat mesenchymal stem cells have been genetically engineered using *ex vivo* retroviral transduction to overexpress the prosurvival gene Akt1, a protein kinase that is the product of an oncogene (Mangi et al., 2003).

[0291] The exogenous proteins expressed by genetically altered mesenchymal stem cells can exert their therapeutic effect by targeting cells present in the environment of the stem cell. The altered stem cells can also respond to the exogenous protein that it secretes. The ability of these altered cells to respond to the transgene, as well as other local signals *in vivo* enables the development of gene therapy strategies that exploit the benefits of particular therapeutic proteins. For example, mesenchymal stem cells that express bone-promoting osteogenic factors can express factors that promote bone growth and can also respond, differentiate, and participate in the bone formation process (Turgeman et al., 2002). Mesenchymal stem cells have also been used to repair defects in cartilage and ligaments (Pend and Huard, 2003), and express a furin-cleavable insulin gene (Sasaki et al., 2003).

Therapeutic Compositions

[0292] The invention further provides agents identified using a screening assay of the invention, and compositions comprising the agents, subject polypeptides, subject polynucleotides, recombinant vectors, and/or host cells, including pharmaceutical compositions for therapeutic administration. The subject compositions can be formulated using well-known reagents and methods. These compositions can include a buffer, which is selected according to the desired use of the agent, polypeptide, polynucleotide, recombinant vector, or host cell, and can also include other substances appropriate to the intended use. Those skilled in the art can readily select an appropriate buffer, a wide variety of which are known in the art, suitable for an intended use.

1. Excipients and Formulations

[0293] In some embodiments, compositions are provided in formulation with pharmaceutically acceptable excipients, a wide variety of which are known in the art (Gennaro, 2000; Ansel et al., 1999; Kibbe et al., 2000). Pharmaceutically acceptable excipients, such as vehicles, adjuvants, carriers or diluents, are readily available to the public. Moreover, pharmaceutically acceptable auxiliary substances, such as pH adjusting and buffering agents, tonicity adjusting agents, stabilizers, wetting agents and the like, are readily available to the public.

[0294] In pharmaceutical dosage forms, the compositions of the invention can be administered in the form of their pharmaceutically acceptable salts, or they can also be used alone or in appropriate association, as well as in combination, with other pharmaceutically active compounds. The subject compositions are formulated in accordance to the mode of potential administration. Administration of the agents can be achieved in various ways, including oral, buccal, nasal, rectal, parenteral, intraperitoneal, intradermal, transdermal, subcutaneous, intravenous, intra-arterial, intracardiac, intraventricular, intracranial, intratracheal, and intrathecal administration, etc., or otherwise by implantation or inhalation. Thus, the subject compositions can be formulated into preparations in solid, semi-solid, liquid or gaseous forms, such as tablets, capsules, powders, granules, ointments, solutions, suppositories, injections, inhalants and aerosols. The following methods and excipients are merely exemplary and are in no way limiting.

[0295] For oral preparations, the agents, polynucleotides, and polypeptides can be used alone or in combination with appropriate additives to make tablets, powders, granules or capsules, for example, with conventional additives, such as lactose, mannitol, corn starch, or potato starch; with binders, such as crystalline cellulose, cellulose derivatives, acacia, corn starch, or gelatins; with disintegrators, such as corn starch, potato starch, or sodium carboxymethylcellulose; with lubricants, such as talc or magnesium stearate; and if desired, with diluents, buffering agents, moistening agents, preservatives, and flavoring agents.

[0296] Suitable excipient vehicles are, for example, water, saline, dextrose, glycerol, ethanol, or the like, and combinations thereof. In addition, if desired, the vehicle can contain minor amounts of auxiliary substances such as wetting or emulsifying agents or pH buffering agents. Actual methods of preparing such dosage forms are known, or will be apparent, to those skilled in the art (Remington, 1985). The composition or formulation to be

administered will, in any event, contain a quantity of the agent adequate to achieve the desired state in the subject being treated.

[0297] The agents, polynucleotides, and polypeptides can be formulated into preparations for injection by dissolving, suspending or emulsifying them in an aqueous or nonaqueous solvent, such as vegetable or other similar oils, synthetic aliphatic acid glycerides, esters of higher aliphatic acids or propylene glycol; and if desired, with conventional additives such as solubilizers, isotonic agents, suspending agents, emulsifying agents, stabilizers and preservatives. Other formulations for oral or parenteral delivery can also be used, as conventional in the art

[0298] The agents, polynucleotides, and polypeptides can be utilized in aerosol formulation to be administered via inhalation. The compounds of the present invention can be formulated into pressurized acceptable propellants such as dichlorodifluoromethane, propane, nitrogen, and the like. Further, the agent, polynucleotides, or polypeptide composition may be converted to powder form for administration intranasally or by inhalation, as conventional in the art.

[0299] Furthermore, the agents can be made into suppositories by mixing with a variety of bases such as emulsifying bases or water-soluble bases. The compounds of the present invention can be administered rectally via a suppository. The suppository can include vehicles such as cocoa butter, carbowaxes and polyethylene glycols, which melt at body temperature, yet are solidified at room temperature.

[0300] A polynucleotide, polypeptide, or other modulator, can also be introduced into tissues or host cells by other routes, such as viral infection, microinjection, or vesicle fusion. For example, expression vectors can be used to introduce nucleic acid compositions into a cell as described above. Further, jet injection can be used for intramuscular administration (Furth et al., 1992). The DNA can be coated onto gold microparticles, and delivered intradermally by a particle bombardment device, or "gene gun" as described in the literature (Tang et al., 1992), where gold microprojectiles are coated with the DNA, then bombarded into skin cells.

[0301] Unit dosage forms for oral or rectal administration such as syrups, elixirs, and suspensions can be provided wherein each dosage unit, for example, teaspoonful, tablespoonful, tablet, or suppository, contains a predetermined amount of the composition containing one or more agents. Similarly, unit dosage forms for injection or intravenous

administration can comprise the agent(s) in a composition as a solution in sterile water, normal saline or another pharmaceutically acceptable carrier.

2. Active Agents (or Modulators)

[0302] The nucleic acid, polypeptide, and modulator compositions of the subject invention find use as therapeutic agents in situations where one wishes to modulate an activity of a subject polypeptide in a host, particularly the activity of the subject polypeptides, or to provide or inhibit the activity at a particular anatomical site. Thus, the compositions are useful in treating disorders associated with an activity of a subject polypeptide. The following provides further details of active agents of the present invention.

a) Antisense Oligonucleotides

[0303] In certain embodiments of the invention, the active agent is an agent that modulates, and generally decreases or down regulates, the expression of a gene encoding a target protein in a host, *i.e.*, antisense molecules. Anti-sense reagents include antisense oligonucleotides (ODN), *i.e.*, synthetic ODN having chemical modifications from native nucleic acids, or nucleic acid constructs that express such anti-sense molecules as RNA. The antisense sequence is complementary to the mRNA of the targeted gene, and inhibits expression of the targeted gene products. Antisense molecules inhibit gene expression through various mechanisms, *e.g.*, by reducing the amount of mRNA available for translation, through activation of RNase H, or steric hindrance. One or a combination of antisense molecules can be administered, where a combination can comprise multiple different sequences.

[0304] Antisense molecules can be produced by expression of all or a part of the target gene sequence in an appropriate vector, where the transcriptional initiation is oriented such that an antisense strand is produced as an RNA molecule. Alternatively, the antisense molecule is a synthetic oligonucleotide. Antisense oligonucleotides can be chemically synthesized by methods known in the art (Wagner et al., 1993; Milligan et al., 1993). Oligonucleotides can be chemically modified from the native phosphodiester structure to increase their intracellular stability and binding affinity, for example, as described in detail above. Antisense oligonucleotides will generally be at least about 7, at least about 12, or at least about 20 nucleotides in length, and not more than about 500, not more than about 50, or not more than about 35 nucleotides in length, where the length is governed by efficiency of inhibition, and specificity, including absence of cross-reactivity, and the like. Short

oligonucleotides, of from about 7 to about 8 bases in length, can be strong and selective inhibitors of gene expression (Wagner et al., 1996).

[0305] A specific region or regions of the endogenous sense strand mRNA sequence is chosen to be complemented by the antisense sequence. Selection of a specific sequence for the oligonucleotide can use an empirical method, where several candidate sequences are assayed for inhibition of expression of the target gene in an *in vitro* or animal model. A combination of sequences can also be used, where several regions of the mRNA sequence are selected for antisense complementation.

[0306] As an alternative to anti-sense inhibitors, catalytic nucleic acid compounds, *e.g.*, ribozymes, or anti-sense conjugates can be used to inhibit gene expression. Ribozymes can be synthesized *in vitro* and administered to the patient, or can be encoded in an expression vector, from which the ribozyme is synthesized in the targeted cell (WO 9523225; Beigelman et al., 1995). Examples of oligonucleotides with catalytic activity are described in WO 9506764. Conjugates of anti-sense ODN with a metal complex, *e.g.*, terpyridyl Cu(II), capable of mediating mRNA hydrolysis are described in Bashkin *et al.*, 1995.

b) Interfering RNA

[0307] In some embodiments, the active agent is an interfering RNA (RNAi), including dsRNAi. RNA interference provides a method of silencing eukaryotic genes. Double stranded RNA can induce the homology-dependent degradation of its cognate mRNA in *C. elegans*, fungi, plants, *Drosophila*, and mammals (Gaudilliere et al., 2002). Use of RNAi to reduce a level of a particular mRNA and/or protein is based on the interfering properties of double-stranded RNA derived from the coding regions of a gene. The technique reduces the time between identifying an interesting gene sequence and understanding its function, and thus is an efficient high-throughput method for disrupting gene function (O'Neil, 2001). RNAi can also help identify the biochemical mode of action of a drug and to identify other genes encoding products that can respond or interact with specific compounds.

[0308] In one embodiment of the invention, complementary sense and antisense RNAs derived from a substantial portion of the subject polynucleotide are synthesized *in vitro*. The resulting sense and antisense RNAs are annealed in an injection buffer, and the double-stranded RNA injected or otherwise introduced into the subject, *i.e.*, in food or by immersion in buffer containing the RNA (Gaudilliere et al., 2002; O'Neil et al., 2001; WO99/32619). In another embodiment, dsRNA derived from a gene of the present invention is generated *in vivo* by simultaneously expressing both sense and antisense RNA from

appropriately positioned promoters operably linked to coding sequences in both sense and antisense orientations.

c) Peptides and Modified Peptides

[0309] In some embodiments of the present invention, the active agent is a peptide. Suitable peptides include peptides of from about 3 amino acids to about 50, from about 5 to about 30, or from about 10 to about 25 amino acids in length. In some embodiments, a peptide has a sequence of from about 3 amino acids to about 50, from about 5 to about 30, or from about 10 to about 25 amino acids of corresponding naturally-occurring protein. In some embodiments, a peptide exhibits one or more of the following activities: inhibits binding of a subject polypeptide to an interacting protein or other molecule; inhibits subject polypeptide binding to a second polypeptide molecule; inhibits a signal transduction activity of a subject polypeptide; inhibits an enzymatic activity of a subject polypeptide; or inhibits a DNA binding activity of a subject polypeptide.

[0310] Peptides can include naturally-occurring and non-naturally occurring amino acids. Peptides can comprise D-amino acids, a combination of D- and L-amino acids, and various "designer" amino acids (*e.g.*, β -methyl amino acids, C α -methyl amino acids, and N α -methyl amino acids, etc.) to convey special properties. Additionally, peptides can be cyclic. Peptides can include non-classical amino acids in order to introduce particular conformational motifs. Any known non-classical amino acid can be used. Non-classical amino acids include, but are not limited to, 1,2,3,4-tetrahydroisoquinoline-3-carboxylate; (2S,3S)-methylphenylalanine, (2S,3R)-methyl-phenylalanine, (2R,3S)-methyl-phenylalanine and (2R,3R)-methyl-phenylalanine; 2-aminotetrahydronaphthalene-2-carboxylic acid; hydroxy-1,2,3,4-tetrahydroisoquinoline-3-carboxylate; β -carboline (D and L); HIC (histidine isoquinoline carboxylic acid); and HIC (histidine cyclic urea). Amino acid analogs and peptidomimetics can be incorporated into a peptide to induce or favor specific secondary structures, including, but not limited to, LL-Acp (LL-3-amino-2-propenidone-6-carboxylic acid), a β -turn inducing dipeptide analog; β -sheet inducing analogs; β -turn inducing analogs; α -helix inducing analogs; γ -turn inducing analogs; Gly-Ala turn analogs; amide bond isostere; or tretrazol, and the like.

[0311] A peptide can be a depsipeptide, which can be linear or cyclic (Kuisle et al., 1999). Linear depsipeptides can comprise rings formed through S-S bridges, or through an hydroxy or a mercapto group of an hydroxy-, or mercapto-amino acid and the carboxyl group of another amino- or hydroxy-acid but do not comprise rings formed only through peptide or

ester links derived from hydroxy carboxylic acids. Cyclic depsipeptides contain at least one ring formed only through peptide or ester links, derived from hydroxy carboxylic acids.

[0312] Peptides can be cyclic or bicyclic. For example, the C-terminal carboxyl group or a C-terminal ester can be induced to cyclize by internal displacement of the -OH or the ester (-OR) of the carboxyl group or ester respectively with the N-terminal amino group to form a cyclic peptide. For example, after synthesis and cleavage to give the peptide acid, the free acid is converted to an activated ester by an appropriate carboxyl group activator such as dicyclohexylcarbodiimide (DCC) in solution, for example, in methylene chloride (CH_2Cl_2), dimethyl formamide (DMF) mixtures. The cyclic peptide is then formed by internal displacement of the activated ester with the N-terminal amine. Internal cyclization as opposed to polymerization can be enhanced by use of very dilute solutions. Methods for making cyclic peptides are well known in the art.

[0313] The term "bicyclic" refers to a peptide with two ring closures formed by covalent linkages between amino acids. A covalent linkage between two nonadjacent amino acids constitutes a ring closure, as does a second covalent linkage between a pair of adjacent amino acids which are already linked by a covalent peptide linkage. The covalent linkages forming the ring closures can be amide linkages, *i.e.*, the linkage formed between a free amino on one amino acid and a free carboxyl of a second amino acid, or linkages formed between the side chains or "R" groups of amino acids in the peptides. Thus, bicyclic peptides can be "true" bicyclic peptides, *i.e.*, peptides cyclized by the formation of a peptide bond between the N-terminus and the C-terminus of the peptide, or they can be "depsi-bicyclic" peptides, *i.e.*, peptides in which the terminal amino acids are covalently linked through their side chain moieties.

[0314] A desamino or decarboxy residue can be incorporated at the terminal ends of the peptide, so that there is no terminal amino or carboxyl group, to decrease susceptibility to proteases or to restrict conformation. C-terminal functional groups include amide, amide lower alkyl, amide di (lower alkyl), lower alkoxy, hydroxy, and carboxy, and the lower ester derivatives thereof, and the pharmaceutically acceptable salts thereof.

[0315] In addition to the foregoing N-terminal and C-terminal modifications, a peptide or peptidomimetic can be modified with or covalently coupled to one or more of a variety of hydrophilic polymers to increase solubility and circulation half-life of the peptide. Suitable nonproteinaceous hydrophilic polymers for coupling to a peptide include, but are not limited to, polyalkylethers as exemplified by polyethylene glycol and polypropylene glycol,

polylactic acid, polyglycolic acid, polyoxyalkenes, polyvinylalcohol, polyvinylpyrrolidone, cellulose and cellulose derivatives, dextran, and dextran derivatives. Generally, such hydrophilic polymers have an average molecular weight ranging from about 500 to about 100,000 daltons, from about 2,000 to about 40,000 daltons, or from about 5,000 to about 20,000 daltons. The peptide can be derivatized with or coupled to such polymers using any of the methods set forth in Zallipsky, 1995; Monfardini et al., 1995; U.S. Pat. Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192; 4,179,337, or WO 95/34326.

d) Antibodies

[0316] The invention provides antibodies that specifically recognize a particular polypeptide. Antibodies are obtained by immunizing a host animal with peptides, polynucleotides encoding polypeptides, or cells, each comprising all or a portion of the target protein ("immunogen"). Suitable host animals include rodents (*e.g.*, mouse, rat, guinea pig, hamster), cattle (*e.g.*, sheep, pig, cow, horse, goat), cat, dog, chicken, primate, monkey, and rabbit. The origin of the protein immunogen can be any species, including mouse, human, rat, monkey, avian, insect, reptile, or crustacean. The host animal will generally be a different species than the immunogen, *e.g.*, a human protein used to immunize mice. Methods of antibody production are well known in the art (Howard and Bethell, 2000; Harlow et al., 1998; Harlow and Lane, 1988).

[0317] The immunogen can comprise the complete protein, or fragments and derivatives thereof, or proteins expressed on cell surfaces. Immunogens comprise all or a part of one of the subject proteins, where these amino acids contain post-translational modifications, such as glycosylation, found on the native target protein. Immunogens comprising protein extracellular domains are produced in a variety of ways known in the art, *e.g.*, expression of cloned genes using conventional recombinant methods, or isolation from tumor cell culture supernatants, etc. The immunogen can also be expressed *in vivo* from a polynucleotide encoding the immunogenic peptide introduced into the host animal.

[0318] Polyclonal antibodies are prepared by conventional techniques. These include immunizing the host animal *in vivo* with the target protein (or immunogen) in substantially pure form, for example, comprising less than about 1% contaminant. The immunogen can comprise the complete target protein, fragments, or derivatives thereof. To increase the immune response of the host animal, the target protein can be combined with an adjuvant; suitable adjuvants include alum, dextran, sulfate, large polymeric anions, and oil & water emulsions, *e.g.*, Freund's adjuvant (complete or incomplete). The target protein can also be

conjugated to synthetic carrier proteins or synthetic antigens. The target protein is administered to the host, usually intradermally, with an initial dosage followed by one or more, usually at least two, additional booster dosages. Following immunization, blood from the host will be collected, followed by separation of the serum from blood cells. The immunoglobulin present in the resultant antiserum can be further fractionated using known methods, such as ammonium salt fractionation, or DEAE chromatography and the like.

[0319] The method of producing polyclonal antibodies can be varied in some embodiments of the present invention. For example, instead of using a single substantially isolated polypeptide as an immunogen, one may inject a number of different immunogens into one animal for simultaneous production of a variety of antibodies. In addition to protein immunogens, the immunogens can be nucleic acids (*e.g.*, in the form of plasmids or vectors) that encode the proteins, with facilitating agents, such as liposomes, microspheres, etc, or without such agents, such as "naked" DNA.

[0320] Antibodies can also be prepared using a library approach. Briefly, mRNA is extracted from the spleens of immunized animals to isolate antibody-encoding sequences. The extracted mRNA may be used to make cDNA libraries. Such a cDNA library may be normalized and subtracted in a manner conventional in the art, for example, to subtract out cDNA hybridizing to mRNA of non-immunized animals. The remaining cDNA may be used to create proteins and for selection of antibody molecules or fragments that specifically bind to the immunogen. The cDNA clones of interest, or fragments thereof, can be introduced into an *in vitro* expression system to produce the desired antibodies, as described herein.

[0321] In a further embodiment, polyclonal antibodies can be prepared using phage display libraries, conventional in the art. In this method, a collection of bacteriophages displaying antibody properties on their surfaces are made to contact subject polypeptides, or fragments thereof. Bacteriophages displaying antibody properties that specifically recognize the subject polypeptides are selected, amplified, for example, in *E. coli*, and harvested. Such a method typically produces single chain antibodies

[0322] Monoclonal antibodies are also produced by conventional techniques, such as fusing an antibody-producing plasma cell with an immortal cell to produce hybridomas. Suitable animals will be used, *e.g.*, to raise antibodies against a mouse polypeptide of the invention, the host animal will generally be a hamster, guinea pig, goat, chicken, or rabbit, and the like. Generally, the spleen and/or lymph nodes of an immunized host animal provide the source of plasma cells, which are immortalized by fusion with myeloma cells to produce

hybridoma cells. Culture supernatants from individual hybridomas are screened using standard techniques to identify clones producing antibodies with the desired specificity. The antibody can be purified from the hybridoma cell supernatants or from ascites fluid present in the host by conventional techniques, *e.g.*, affinity chromatography using antigen, *e.g.*, the subject protein, bound to an insoluble support, *i.e.*, protein A sepharose, etc.

[0323] The antibody can be produced as a single chain, instead of the normal multimeric structure of the immunoglobulin molecule. Single chain antibodies have been previously described (*i.e.*, Jost et al., 1994). DNA sequences encoding parts of the immunoglobulin, for example, the variable region of the heavy chain and the variable region of the light chain are ligated to a spacer, such as one encoding at least about four small neutral amino acids, *i.e.*, glycine or serine. The protein encoded by this fusion allows the assembly of a functional variable region that retains the specificity and affinity of the original antibody.

[0324] The invention also provides intrabodies that are intracellularly expressed single-chain antibody molecules designed to specifically bind and inactivate target molecules inside cells. Intrabodies have been used in cell assays and in whole organisms (Chen et al., 1994; Hassanzadeh et al., 1998). Inducible expression vectors can be constructed with intrabodies that react specifically with a protein of the invention. These vectors can be introduced into host cells and model organisms.

[0325] The invention also provides "artificial" antibodies, *e.g.*, antibodies and antibody fragments produced and selected *in vitro*. In some embodiments, these antibodies are displayed on the surface of a bacteriophage or other viral particle, as described above. In other embodiments, artificial antibodies are present as fusion proteins with a viral or bacteriophage structural protein, including, but not limited to, M13 gene III protein. Methods of producing such artificial antibodies are well known in the art (U.S. Patent Nos. 5,516,637; 5,223,409; 5,658,727; 5,667,988; 5,498,538; 5,403,484; 5,571,698; and 5,625,033). The artificial antibodies, selected for example, on the basis of phage binding to selected antigens, can be fused to a Fc fragment of an immunoglobulin for use as a therapeutic, as described, for example, in US 5,116,964 or WO 99/61630. Antibodies of the invention can be used to modulate biological activity of cells, either directly or indirectly. A subject antibody can modulate the activity of a target cell, with which it has primary interaction, or it can modulate the activity of other cells by exerting secondary effects, *i.e.*, when the primary targets interact or communicate with other cells. The antibodies of the invention can be administered to

mammals, and the present invention includes such administration, particularly for therapeutic and/or diagnostic purposes in humans.

[0326] Antibodies may be administered by injection systemically, such as by intravenous injection; or by injection or application to the relevant site, such as by direct injection into a tumor, or direct application to the site when the site is exposed in surgery; or by topical application, such as if the disorder is on the skin, for example.

[0327] For *in vivo* use, particularly for injection into humans, in some embodiments it is desirable to decrease the antigenicity of the antibody. An immune response of a recipient against the antibody may potentially decrease the period of time that the therapy is effective. Methods of humanizing antibodies are known in the art. The humanized antibody can be the product of an animal having transgenic human immunoglobulin genes, *e.g.*, constant region genes (*e.g.*, Grosveld and Kolias, 1992; Murphy and Carter, 1993; Pinkert, 1994; and International Patent Applications WO 90/10077 and WO 90/04036). Alternatively, the antibody of interest can be engineered by recombinant DNA techniques to substitute the CH1, CH2, CH3, hinge domains, and/or the framework domain with the corresponding human sequence (see, *e.g.*, WO 92/02190). Both polyclonal and monoclonal antibodies made in non-human animals may be "humanized" before administration to human subjects.

[0328] Chimeric immunoglobulin genes constructed with immunoglobulin cDNA are known in the art (Liu et al. 1987a; Liu et al. 1987b). Messenger RNA is isolated from a hybridoma or other cell producing the antibody and used to produce cDNA. The cDNA of interest can be amplified by the polymerase chain reaction using specific primers (U.S. Patent nos. 4,683,195 and 4,683,202). Alternatively, a library is made and screened to isolate the sequence of interest. The DNA sequence encoding the variable region of the antibody is then fused to human constant region sequences. The sequences of human constant regions genes are known in the art (Kabat et al., 1991). Human C region genes are readily available from known clones. The choice of isotype will be guided by the desired effector functions, such as complement fixation, or antibody-dependent cellular cytotoxicity. IgG1, IgG3 and IgG4 isotypes, and either of the kappa or lambda human light chain constant regions can be used. The chimeric, humanized antibody is then expressed by conventional methods.

[0329] Consensus sequences of heavy ("H") and light ("L") J regions can be used to design oligonucleotides for use as primers to introduce useful restriction sites into the J region for subsequent linkage of V region segments to human C region segments. C region

cDNA can be modified by site directed mutagenesis to place a restriction site at the analogous position in the human sequence.

[0330] A convenient expression vector for producing antibodies is one that encodes a functionally complete human CH or CL immunoglobulin sequence, with appropriate restriction sites engineered so that any VH or VL sequence can be easily inserted and expressed, such as plasmids, retroviruses, YACs, or EBV derived episomes, and the like. In such vectors, splicing usually occurs between the splice donor site in the inserted J region and the splice acceptor site preceding the human C region, and also at the splice regions that occur within the human CH exons. Polyadenylation and transcription termination occur at native chromosomal sites downstream of the coding regions. The resulting chimeric antibody can be joined to any strong promoter, including retroviral LTRs, *e.g.*, SV-40 early promoter, (Okayama, et al. 1983), Rous sarcoma virus LTR (Gorman et al. 1982), and Moloney murine leukemia virus LTR (Grosschedl et al. 1985), or native immunoglobulin promoters.

[0331] In yet other embodiments, the antibodies can be fully human antibodies. For example, xenogenic antibodies, which are produced in animals that are transgenic for human antibody genes, can be employed. By xenogenic human antibodies is meant antibodies that are fully human antibodies, with the exception that they are produced in a non-human host that has been genetically engineered to express human antibodies. (*e.g.*, WO 98/50433; WO 98,24893 and WO 99/53049).

[0332] Antibody fragments, such as Fv, F(ab')₂ and Fab can be prepared by cleavage of the intact protein, *e.g.*, by protease or chemical cleavage. These fragments can include heavy and light chain variable regions. Alternatively, a truncated gene can be designed, *e.g.*, a chimeric gene encoding a portion of the F(ab')₂ fragment that includes DNA sequences encoding the CH1 domain and hinge region of the H chain, followed by a translational stop codon. The antibodies of the present invention may be administered alone or in combination with other molecules for use as a therapeutic, for example, by linking the antibody to cytotoxic agent, as discussed above, or to a radioactive molecule. Radioactive antibodies that are specific to a cancer cell, disease cell, or virus-infected cell may be able to deliver a sufficient dose of radioactivity to kill such cancer cell, disease cell, or virus-infected cell. The antibodies of the present invention can also be used in assays for detection of the subject polypeptides. In some embodiments, the assay is a binding assay that detects binding of a polypeptide with an antibody specific for the polypeptide; the subject polypeptide or antibody can be immobilized, while the subject polypeptide and/or antibody can be detectably-labeled.

For example, the antibody can be directly labeled or detected with a labeled secondary antibody. That is, suitable, detectable labels for antibodies include direct labels, which label the antibody to the protein of interest, and indirect labels, which label an antibody that recognizes the antibody to the protein of interest.

[0333] These labels include radioisotopes, including, but not limited to ^{64}Cu , ^{67}Cu , ^{90}Y , ^{124}I , ^{125}I , ^{131}I , ^{137}Cs , ^{186}Re , ^{211}At , ^{212}Bi , ^{213}Bi , ^{223}Ra , ^{241}Am , and ^{244}Cm ; enzymes having detectable products (*e.g.*, luciferase, β -galactosidase, and the like); fluorescers and fluorescent labels, *e.g.*, as provided herein; fluorescence emitting metals, *e.g.*, ^{152}Eu , or others of the lanthanide series, attached to the antibody through metal chelating groups such as EDTA; chemiluminescent compounds, *e.g.*, luminol, isoluminol, or acridinium salts; and bioluminescent compounds, *e.g.*, luciferin, or aequorin (green fluorescent protein), specific binding molecules, *e.g.*, magnetic particles, microspheres, nanospheres, and the like.

[0334] Alternatively, specific-binding pairs may be used, involving, *e.g.*, a second stage antibody or reagent that is detectably-labeled and that can amplify the signal. For example, a primary antibody can be conjugated to biotin, and horseradish peroxidase-conjugated streptavidin added as a second stage reagent. Digoxin and antidigoxin provide another such pair. In other embodiments, the secondary antibody can be conjugated to an enzyme such as peroxidase in combination with a substrate that undergoes a color change in the presence of the peroxidase. The absence or presence of antibody binding can be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, or scintillation counting. Such reagents and their methods of use are well known in the art.

e) Peptide Aptamers

[0335] Another suitable agent for modulating an activity of a subject polypeptide is a peptide aptamer. Peptide aptamers are peptides or small polypeptides that act as dominant inhibitors of protein function. Peptide aptamers specifically bind to target proteins, blocking their functional ability (Kolonin and Finley, 1998). Due to the highly selective nature of peptide aptamers, they can be used not only to target a specific protein, but also to target specific functions of a given protein (*e.g.*, a signaling function). Further, peptide aptamers can be expressed in a controlled fashion by use of promoters which regulate expression in a temporal, spatial or inducible manner. Peptide aptamers act dominantly, therefore, they can be used to analyze proteins for which loss-of-function mutants are not available.

[0336] Peptide aptamers that bind with high affinity and specificity to a target protein can be isolated by a variety of techniques known in the art. Peptide aptamers can be isolated from random peptide libraries by yeast two-hybrid screens (Xu et al., 1997). They can also be isolated from phage libraries (Hoogenboom et al., 1998) or chemically generated peptides/libraries.

Therapeutic Applications

[0337] The instant invention provides various therapeutic methods. In some embodiments, methods of modulating, including increasing and inhibiting, a biological activity of a subject protein are provided. In some embodiments, methods of modulating an enzymatic activity of a subject protein are provided. In some embodiments, methods of increasing the level of enzymatically active subject protein are provided, while in some embodiments, methods of decreasing a level of enzymatically active subject protein are provided.

[0338] In some embodiments, methods of modulating enzymatic activity of a subject protein are provided. In other embodiments, methods of modulating a signal transduction activity of a subject protein are provided. In further embodiments, methods of modulating interaction of a subject protein with another, interacting protein or other macromolecule (*e.g.*, DNA, carbohydrate, lipid) are provided. In further embodiments, methods of modulating transport activity of a subject protein are provided. In further embodiments, methods of modulating phospholipase activity of a subject protein are provided. In further embodiments, methods of modulating polymerase activity of a subject protein are provided. In further embodiments, methods of modulating nuclease activity of a subject protein are provided.

[0339] As mentioned above, an effective amount of the active agent (*e.g.*, small molecule, antibody specific for a subject polypeptide, a subject polypeptide, or a subject polynucleotide) is administered to the host, where "effective amount" means a dosage sufficient to produce a desired effect or result. In some embodiments, the desired result is at least a reduction in a given biological activity of a subject polypeptide as compared to a control, for example, a decreased level of enzymatically active subject protein in the individual, or in a localized anatomical site in the individual. In further embodiments, the desired result is at least an increase in a biological activity of a subject polypeptide as compared to a control, for example an increased level of enzymatically active subject protein in the individual, or in a localized anatomical site in the individual.

[0340] Typically, the compositions of the instant invention will contain from less than about 1% to about 95% of the active ingredient, about 10% to about 50%. Generally, between about 100 mg and about 500 mg will be administered to a child and between about 500 mg and about 5 grams will be administered to an adult.

[0341] Other effective dosages can be readily determined by one of ordinary skill in the art through routine trials establishing dose response curves, for example, the amount of agent necessary to increase a level of active subject polypeptide can be calculated from *in vitro* experimentation. Those of skill will readily appreciate that dose levels can vary as a function of the specific compound, the severity of the symptoms, and the susceptibility of the subject to side effects, and preferred dosages for a given compound are readily determinable by those of skill in the art by a variety of means. For example, in order to calculate the polypeptide, polynucleotide, or modulator dose, those skilled in the art can use readily available information with respect to the amount necessary to have the desired effect, depending upon the particular agent used.

[0342] The active agent(s) can be administered to the host via any convenient means capable of resulting in the desired result. Administration is generally by injection and often by injection to a localized area. The frequency of administration will be determined by the care given based on patient responsiveness. For example, the agents may be administered daily, weekly, or as conventionally determined appropriate.

[0343] A variety of hosts are treatable according to the subject methods. The host, or patient, may be from any animal species, and will generally be mammalian, *e.g.*, primate sp., *e.g.*, monkeys, chimpanzees, and particularly humans; rodents, including mice, rats and hamsters, guinea pig; rabbits; cattle, including equines, bovines, pig, sheep, goat, canines; felines; etc. Animal models are of interest for experimental investigations, providing a model for treatment of human disease.

1. Proliferative Conditions

[0344] In some embodiments, a protein of the present invention is involved in the control of cell proliferation, and an agent of the invention inhibits undesirable cell proliferation. Such agents are useful for treating disorders that involve abnormal cell proliferation, including, but not limited to, cancer, psoriasis, and scleroderma. Whether a particular agent and/or therapeutic regimen of the invention is effective in reducing unwanted cellular proliferation, *e.g.*, in the context of treating cancer, can be determined using standard methods. For example, the number of cancer cells in a biological sample (*e.g.*, blood, a

biopsy sample, and the like), can be determined. The tumor mass can be determined using standard radiological or biochemical methods.

[0345] Tumors that can be treated using the methods of the instant invention include carcinomas, *e.g.*, colorectal, prostate, breast, bone, kidney, skin, melanoma, ductal, endometrial, stomach or other organ of the gastrointestinal tract, pancreatic, mesothelioma, dysplastic oral mucosa, invasive oral cancer, tracheal cancer, non-small cell lung carcinoma ("NSCL"), transitional and squamous cell urinary carcinoma; brain cancer and neurological malignancies, *e.g.*, neuroblastoma, glioblastoma, astrocytoma, and gliomas; lymphomas and leukemias such as myeloid leukemia, myelogenous leukemia, hematological malignancies, such as childhood acute leukemia, non-Hodgkin's lymphomas, chronic lymphocytic leukemia, malignant cutaneous T-cell lymphoma, mycosis fungoides, non-MF cutaneous T-cell lymphoma, lymphomatoid papulosis, T-cell rich cutaneous lymphoid hyperplasia, bullous pemphigoid, discoid lupus erythematosus, lichen planus, and human follicular lymphoma; cancers of the reproductive system, *e.g.*, cervical and ovarian cancers and testicular cancers; liver cancers including hepatocellular carcinoma ("HCC") and tumors of the biliary duct; multiple myelomas; tumors of the esophageal tract; other lung cancers and tumors including small cell and clear cell; Hodgkin's lymphomas; adenocarcinoma; and sarcomas, including soft tissue sarcomas.

2. Immunotherapeutic Approaches to Proliferative Conditions

[0346] The polynucleotides, polypeptides, and modulators of the present invention find use in immunotherapy of hyperproliferative disorders, including cancer, neoplastic, and paraneoplastic disorders. That is, the subject molecules can correspond to tumor antigens, of which 1770 have been identified to date (Yu and Restifo, 2002). Immunotherapeutic approaches include passive immunotherapy and vaccine therapy and can accomplish both generic and antigen-specific cancer immunotherapy.

[0347] Passive immunity approaches involve antibodies of the invention that are directed toward specific tumor-associated antigens. Such antibodies can eradicate systemic tumors at multiple sites, without eradicating normal cells. In some embodiments, the antibodies are combined with radioactive components, as provided above, for example, combining the antibody's ability to specifically target tumors with the added lethality of the radioisotope to the tumor DNA.

[0348] Useful antibodies comprise a discrete epitope or a combination of nested epitopes, *i.e.*, a 10-mer epitope and associated peptide multimers incorporating all potential 8-

mers and 9-mers, or overlapping epitopes (Dutoit et al., 2002). Thus a single antibody can interact with one or more epitopes. Further, the antibody can be used alone or in combination with different antibodies, that all recognize either a single or multiple epitopes.

[0349] Neutralizing antibodies can provide therapy for cancer and proliferative disorders. Neutralizing antibodies that specifically recognize a secreted protein or peptide of the invention can bind to the secreted protein or peptide, *e.g.*, in a bodily fluid or the extracellular space, thereby modulating the biological activity of the secreted protein or peptide. For example, neutralizing antibodies specific for secreted proteins or peptides that play a role in stimulating the growth of cancer cells can be useful in modulating the growth of cancer cells. Similarly, neutralizing antibodies specific for secreted proteins or peptides that play a role in the differentiation of cancer cells can be useful in modulating the differentiation of cancer cells.

[0350] Vaccine therapy involves the use of polynucleotides, polypeptides, or agents of the invention as immunogens for tumor antigens (Machiels et al., 2002). For example, peptide-based vaccines of the invention include unmodified subject polypeptides, fragments thereof, and MHC class I and class II-restricted peptide (Knutson et al., 2001), comprising, for example, the disclosed sequences with universal, nonspecific MHC class II-restricted epitopes. Peptide-based vaccines comprising a tumor antigen can be given directly, either alone or in conjunction with other molecules. The vaccines can also be delivered orally by producing the antigens in transgenic plants that can be subsequently ingested (U.S. Patent No. 6,395,964).

[0351] In some embodiments, antibodies themselves can be used as antigens in anti-idiotypic vaccines. That is, administering an antibody to a tumor antigen stimulates B cells to make antibodies to that antibody, which in turn recognize the tumor cells

[0352] Nucleic acid-based vaccines can deliver tumor antigens as polynucleotide constructs encoding the antigen. Vaccines comprising genetic material, such as DNA or RNA, can be given directly, either alone or in conjunction with other molecules. Administration of a vaccine expressing a molecule of the invention, *e.g.*, as plasmid DNA, leads to persistent expression and release of the therapeutic immunogen over a period of time, helping to control unwanted tumor growth.

[0353] In some embodiments, nucleic acid-based vaccines encode subject antibodies. In such embodiments, the vaccines (*e.g.*, DNA vaccines) can include post-transcriptional regulatory elements, such as the post-transcriptional regulatory acting RNA element (WPRE)

derived from Woodchuck Hepatitis Virus. These post-transcriptional regulatory elements can be used to target the antibody, or a fusion protein comprising the antibody and a co-stimulatory molecule, to the tumor microenvironment (Pertl et al., 2003).

[0354] Besides stimulating anti-tumor immune responses by inducing humoral responses, vaccines of the invention can also induce cellular responses, including stimulating T-cells that recognize and kill tumor cells directly. For example, nucleotide-based vaccines of the invention encoding tumor antigens can be used to activate the CD8⁺ cytotoxic T lymphocyte arm of the immune system.

[0355] In some embodiments, the vaccines activate T-cells directly, and in others they enlist antigen-presenting cells to activate T-cells. Killer T-cells are primed, in part, by interacting with antigen-presenting cells, *i.e.*, dendritic cells. In some embodiments, plasmids comprising the nucleic acid molecules of the invention enter antigen-presenting cells, which in turn display the encoded tumor-antigens that contribute to killer T-cell activation. Again, the tumor antigens can be delivered as plasmid DNA constructs, either alone or with other molecules.

[0356] In further embodiments, RNA can be used. For example, dendritic cells can be transfected with RNA encoding tumor antigens (Heiser et al., 2002; Mitchell and Nair, 2000). This approach overcomes the limitations of obtaining sufficient quantities of tumor material, extending therapy to patients otherwise excluded from clinical trials. For example, a subject RNA molecule isolated from tumors can be amplified using RT-PCR. In some embodiments, the RNA molecule of the invention is directly isolated from tumors and transfected into dendritic cells with no intervening cloning steps.

[0357] In some embodiments the molecules of the invention are altered such that the peptide antigens are more highly antigenic than in their native state. These embodiments address the need in the art to overcome the poor *in vivo* immunogenicity of most tumor antigens by enhancing tumor antigen immunogenicity via modification of epitope sequences (Yu and Restifo, 2002).

[0358] Another recognized problem of cancer vaccines is the presence of preexisting neutralizing antibodies. Some embodiments of the present invention overcome this problem by using viral vectors from non-mammalian natural hosts, *i.e.*, avian pox viruses. Alternative embodiments that also circumvent preexisting neutralizing antibodies include genetically engineered influenza viruses, and the use of "naked" plasmid DNA vaccines that contain DNA with no associated protein (Yu and Restifo, 2002).

[0359] All of the immunogenic methods of the invention can be used alone or in combination with other conventional or unconventional therapies. For example, immunogenic molecules can be combined with other molecules that have a variety of antiproliferative effects, or with additional substances that help stimulate the immune response, *i.e.*, adjuvants or cytokines.

[0360] For example, in some embodiments, nucleic acid vaccines encode an alphaviral replicase enzyme, in addition to tumor antigens. This recently discovered approach to vaccine therapy successfully combines therapeutic antigen production with the induction of the apoptotic death of the tumor cell (Yu and Restifo, 2002).

[0361] In certain other embodiments, a DNA or RNA vaccine of the present invention can also be directed against the production of blood vessels in the vicinity of the tumor, a process called antiangiogenesis, thereby depriving the cancer cells of nutrients. For example, the antiangiogenic molecules angiostatin (a fragment of plasminogen), endostatin (a fragment of collagen XVIII), interferon- γ , interferon- γ inducible protein 10, interleukin 12, thrombospondin, platelet factor-4, calreticulin, or its protein fragment vasostatin can be used to treat tumors by suppressing neovascularization and thereby inhibiting growth (Cheng et al., 2001). The antiangiogenesis approach can be used alone, or in conjunction with molecules directed to tumor antigens.

[0362] Furthermore, adjuvants can be used in conjunction with the antibodies and vaccines disclosed herein. Adjuvants help boost the general immune response, for example, concentrating immune cells to the specific area where they are needed. They can be added to a cancer vaccine itself or administered separately, and in some embodiments, a viral vector can be engineered to display adjuvant proteins on its surface.

[0363] Cytokines can also be used to help stimulate immune response. Cytokines act as chemical messengers, recruiting immune cells that help the killer T-cells to the site of attack. An example of a cytokine is granulocyte-macrophage colony-stimulating factor (GM-CSF), which stimulates the proliferation of antigen-presenting cells, thus boosting an organism's response to a cancer vaccine. As with adjuvants, cytokines can be used in conjunction with the antibodies and vaccines disclosed herein. For example, they can be incorporated into the antigen-encoding plasmid or introduced via a separate plasmid, and in some embodiments, a viral vector can be engineered to display cytokines on its surface.

3. Inflammation and Immunity

[0364] In other embodiments, *e.g.*, where the subject polypeptide is involved in modulating inflammation or immune function, the invention provides agents for treating such inflammation or immune disorders. Disease states that are treatable using formulations of the invention include various types of arthritis such as rheumatoid arthritis and osteoarthritis, autoimmune thyroiditis, various chronic inflammatory conditions of the skin, such as psoriasis, the intestine, such as inflammatory bowel disease (IBD), insulin-dependent diabetes, autoimmune diseases such as multiple sclerosis (MS), intestinal immune disorders and systemic lupus erythematosus (SLE), allergic diseases, transplant rejections, adult respiratory distress syndrome, atherosclerosis, ischemic diseases due to closure of the peripheral vasculature, cardiac vasculature, and vasculature in the central nervous system (CNS). After reading the present disclosure, those skilled in the art will recognize other disease states and/or symptoms which might be treated and/or mitigated by the administration of formulations of the present invention.

[0365] Neutralizing antibodies can provide immunosuppressive therapy for inflammatory and autoimmune disorders. Neutralizing antibodies can be used to treat disorders such as, for example, multiple sclerosis, rheumatoid arthritis, inflammatory bowel disease, transplant rejection, and psoriasis. Neutralizing antibodies that specifically recognize a secreted protein or peptide of the invention can bind to the secreted protein or peptide, *e.g.*, in a bodily fluid or the extracellular space, thereby modulating the biological activity of the secreted protein or peptide. For example, neutralizing antibodies specific for secreted proteins or peptides that play a role in activating immune cells are useful as immunosuppressants.

4. Disorders Related to Cell Death

[0366] Where a polypeptide of the invention is involved in modulating cell death, an agent of the invention is useful for treating conditions or disorders relating to cell death (*e.g.*, DNA damage, cell death, apoptosis). Cell death-related indications that can be treated using the methods of the invention to reduce cell death in a eukaryotic cell, include, but are not limited to, cell death associated with Alzheimer's disease, Parkinson's disease, rheumatoid arthritis, autoimmune thyroiditis, septic shock, sepsis, stroke, central nervous system inflammation, intestinal inflammation, osteoporosis, ischemia, reperfusion injury, cardiac muscle cell death associated with cardiovascular disease, polycystic kidney disease, cell death of endothelial cells in cardiovascular disease, degenerative liver disease, multiple

sclerosis, amyotrophic lateral sclerosis, cerebellar degeneration, ischemic injury, cerebral infarction, myocardial infarction, acquired immunodeficiency syndrome (AIDS), myelodysplastic syndromes, aplastic anemia, male pattern baldness, and head injury damage. Also included are conditions in which DNA damage to a cell is induced by external conditions, including but not limited to irradiation, radiomimetic drugs, hypoxic injury, chemical injury, and damage by free radicals. Also included are any hypoxic or anoxic conditions, *e.g.*, conditions relating to or resulting from ischemia, myocardial infarction, cerebral infarction, stroke, bypass heart surgery, organ transplantation, and neuronal damage, etc.

[0367] DNA damage can be detected using any known method, including, but not limited to, a Comet assay (commercially available from Trevigen, Inc.), which is based on alkaline lysis of labile DNA at sites of damage; and immunological assays using antibodies specific for aberrant DNA structures, *e.g.*, 8-OHdG.

[0368] Cell death can be measured using any known method, and is generally measured using any of a variety of known methods for measuring cell viability. Such assays are generally based on entry into the cell of a detectable compound (or a compound that becomes detectable upon interacting with, or being acted on by, an intracellular component) that would normally be excluded from a normal, living cell by its structurally and functionally intact cell membrane. Such compounds include substrates for intracellular enzymes, including, but not limited to, a fluorescent substrate for esterase; dyes that are excluded from living cells, including, but not limited to, trypan blue; and DNA-binding compounds, including, but not limited to, an ethidium compound such as ethidium bromide and ethidium homodimer, and propidium iodide.

[0369] Apoptosis, or programmed cell death, is a regulated process leading to cell death via a series of well-defined morphological changes. Programmed cell death provides a balance for cell growth and multiplication, eliminating unnecessary cells. The default state of the cell is to remain alive. A cell enters the apoptotic pathway when an essential factor is removed from the extracellular environment or when an internal signal is activated. Genes and proteins of the invention that suppress the growth of tumors by activating cell death provide the basis for treatment strategies for hyperproliferative disorders and conditions.

[0370] Apoptosis can be assayed using any known method. Assays can be conducted on cell populations or an individual cell, and include morphological assays and biochemical assays. A non-limiting example of a method of determining the level of apoptosis in a cell

population is TUNEL (TdT-mediated dUTP nick-end labeling) labeling of the 3'-OH free end of DNA fragments produced during apoptosis (Gavrieli et al., 1992). The TUNEL method consists of catalytically adding a nucleotide, which has been conjugated to a chromogen system, a fluorescent tag, or the 3'-OH end of the 180-bp (base pair) oligomer DNA fragments, in order to detect the fragments. The presence of a DNA ladder of 180-bp oligomers is indicative of apoptosis. Procedures to detect cell death based on the TUNEL method are available commercially, *e.g.*, from Boehringer Mannheim (Cell Death Kit) and Oncor (Apoptag Plus).

[0371] Another marker that is currently available is annexin, sold under the trademark APOPTTEST™. This marker is used in the "Apoptosis Detection Kit," which is also commercially available, *e.g.*, from R&D Systems. During apoptosis, a cell membrane's phospholipid asymmetry changes such that the phospholipids are exposed on the outer membrane. Annexins are a homologous group of proteins that bind phospholipids in the presence of calcium. A second reagent, propidium iodide (PI), is a DNA binding fluorochrome. When a cell population is exposed to both reagents, apoptotic cells stain positive for annexin and negative for PI, necrotic cells stain positive for both, live cells stain negative for both. Other methods of testing for apoptosis are known in the art and can be used, including, *e.g.*, the method disclosed in U.S. Patent No. 6,048,703.

5. Other Pathological Conditions

[0372] Other pathological conditions that can be treated using the methods of the instant invention include disorders of hematopoiesis, cell differentiation, disorders of ion channels, *e.g.*, cystic fibrosis, and tissue or organ hypertrophy, viral disorders, including acquired immunodeficiency syndrome (AIDS), angiogenesis, metastasis, metabolic disorders such as diabetes and obesity, cardiovascular disorders such as congestive heart failure and stroke, male erectile dysfunction, and the disorders described throughout the specification.

Investigative Applications

[0373] The subject nucleic acid compositions find use in a variety of different investigative applications. Applications of interest include identifying genomic DNA sequence using molecules of the invention, identifying homologs of molecules of the invention, creating a source of novel promoter elements, identifying expression regulatory factors, creating a source of probes and primers for hybridization applications, identifying expression patterns in biological specimens; preparing cell or animal models to investigate

the function of the molecules of the invention, and preparing *in vitro* models to investigate the function of the molecules of the invention.

Genomic DNA Sequences

[0374] Human genomic polynucleotide sequences corresponding to molecules of the present invention are identified by conventional means, such as, for example, by probing a genomic DNA library with all or a portion of the polynucleotide sequences.

[0375] Homologs are identified by any of a number of methods. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes, as described in detail above. Briefly, a fragment of the provided cDNA can be used as a hybridization probe against a cDNA library from the target organism of interest, under various stringency conditions, *e.g.*, low stringency conditions. The probe can be a large fragment, or one or more short degenerate primers, and is typically labeled. Sequence identity can be determined by hybridization under stringent conditions, as described in detail above. Nucleic acids having a region of substantial identity or sequence similarity to the provided nucleic acid sequences, for example allelic variants, related genes, or genetically altered versions of the gene, bind to the provided sequences under less stringent hybridization conditions.

Promoter Elements and Expression Regulatory Factors

[0376] The sequence of the 5' flanking region can be utilized as promoter elements, including enhancer binding sites that provide for tissue-specific expression and developmental regulation in tissues where the subject genes are expressed, providing promoters that mimic the native pattern of expression. Naturally occurring polymorphisms in the promoter region are useful for determining natural variations in expression, particularly those that may be associated with disease. Promoters or enhancers that regulate the transcription of the polynucleotides of the present invention are obtainable by use of PCR techniques using human tissues, and one or more of the present primers.

[0377] Alternatively, mutations can be introduced into the promoter region to determine the effect of altering expression in experimentally defined systems. Methods for the identification of specific DNA motifs involved in the binding of transcriptional factors are known in the art, for example sequence similarity to known binding motifs, and gel retardation studies (Blackwell et al., 1995; Mortlock et al., 1996; Joulin and Richard-Foy, 1995).

[0378] The regulatory sequences can be used to identify *cis* acting sequences required for transcriptional or translational regulation of expression, especially in different tissues or stages of development, and to identify *cis* acting sequences and *trans*-acting factors that regulate or mediate expression. Such transcription or translational control regions can be operably linked to a gene in order to promote expression of wild type genes or of proteins of interest in cultured cells, embryonic, fetal or adult tissues, and for gene therapy (Hooper, 1993).

Primers and Probes

[0379] Small DNA fragments are useful as primers for reactions that involve nucleic acid hybridization, as described in detail above. Briefly, pairs of primers will be used in amplification reactions, such as PCR. Amplification primers hybridize to complementary strands of DNA, for example, under stringent conditions, and will prime towards each other. In some embodiments a pair of primers will generate an amplification product of at least about 50 nt, or at least about 100 nt. Algorithms for the selection of primer sequences are generally known, and are available in commercial software packages.

[0380] The nucleotides can also be used as probes to identify genomic DNA or gene expression in a biological specimen, as described above and as is well established in the art. Briefly, DNA or mRNA is isolated from a cell sample. Detection of mRNA hybridizing to the subject sequence is indicative of gene expression in the sample. The mRNA can be amplified by RT-PCR, using reverse transcriptase to form a complementary DNA strand, followed by polymerase chain reaction amplification using primers specific for the subject DNA sequences. Alternatively, the mRNA sample is separated by gel electrophoresis, transferred to a suitable support, *e.g.*, nitrocellulose, nylon, *etc.*, and then probed with a fragment of the subject nucleotides as a probe. Other techniques, such as oligonucleotide ligation assays, *in situ* hybridizations, and hybridization to probes arrayed on a solid chip may also find use.

Targeted Mutations for *In Vivo* and *In Vitro* Models

[0381] The sequence of a gene according to the subject invention, including flanking promoter regions and coding regions, can be mutated in various ways known in the art to generate targeted changes, *i.e.*, changes in promoter strength, or sequence of the encoded protein, *etc.* The DNA sequence or protein product of such a mutation will usually be substantially similar to the sequences provided herein. The sequence changes can be

substitutions, insertions, deletions, or a combination thereof. Deletions can further include larger changes, such as deletions of a domain or exon.

[0382] Techniques for *in vitro* mutagenesis of cloned genes are known. Examples of protocols for site specific mutagenesis may be found in Gustin et al., 1993; Barany 1985; Colicelli et al., 1985; Prentki et al., 1984. Methods for site specific mutagenesis can be found in Sambrook et al., 1989 (pp. 15.3-15.108); Weiner et al., 1993; Sayers et al. 1992; Jones and Winistorfer; Barton et al., 1990; Marotti and Tomich 1989; and Zhu, 1989. Such mutated genes can be used to study structure-function relationships of the subject proteins, or to alter properties of the protein that affect its function or regulation. Other modifications of interest include epitope tagging, *e.g.*, with hemagglutinin (HA), FLAG, or *c-myc*. For studies of subcellular localization, fluorescent fusion proteins can be used.

[0383] The subject nucleic acids can be used to generate transgenic, non-human animals and/or site-specific gene modifications in cell lines; suitable methods are known in the art (Grosveld and Kollias, 1992; Hooper, 1993; Murphy and Carter, 1993; Pinkert, 1994). Thus, in some embodiments, the invention provides a non-human transgenic animal comprising, as a transgene integrated into the genome of the animal, a nucleic acid molecule comprising a sequence encoding a subject polypeptide in operable linkage with a promoter, such that the subject polypeptide-encoding nucleic acid molecule is expressed in a cell of the animal. Either a complete or partial sequence of a gene native to the host can be introduced. Alternatively, a complete or partial sequence of a gene exogenous to the host animal, *e.g.*, a human sequence of the subject invention, can be introduced. Transgenic animals can be made through homologous recombination, where the endogenous locus is altered. Thus, DNA constructs for homologous recombination will comprise at least a portion of the human gene or of a gene native to the species of the host animal, wherein the gene has the desired genetic modification(s), and includes regions of homology to the target locus. Methods for generating mammalian cells having targeted gene modifications through homologous recombination are known in the art (Keown et al., 1990).

[0384] Alternatively, a nucleic acid construct is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, and YACs. DNA constructs for random integration need not include regions of homology to mediate recombination.

[0385] Conveniently, markers for positive and negative selection are included. A detectable marker, such as *lac Z* can be introduced into a locus at which up-regulation of expression will result in a detectable change in phenotype.

[0386] Transformed ES or embryonic cells can be used to produce transgenic animals. An embryonic stem (ES) cell line can be a source of embryonic stem cells, or they can be newly obtained from a host animal, *e.g.*, a mouse, rat, or guinea pig. The cells are grown on an appropriate fibroblast-feeder layer or in the presence of leukemia inhibiting factor (LIF). Following transformation, the cells are plated for growth onto a feeder layer in an appropriate medium. Cells containing the relevant construct can be detected by employing a selective medium and analyzing them for the occurrence of homologous recombination or integration of the construct. Positive colonies can be used for embryo manipulation and blastocyst injection. Blastocysts are obtained from 4 to 6 week old super-ovulated females. The ES cells are trypsinized, and the modified cells are injected into the blastocoel of the blastocyst. After injection, the blastocysts are returned to each uterine horn of pseudopregnant female animals that proceed to term. The resulting offspring are screened for the construct. By providing for a different phenotype of the blastocyst and the genetically modified cells, chimeric progeny can be readily detected.

[0387] The chimeric animals are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogeneic or congenic grafts or transplants, or in *in vitro* culture. The transgenic animals can be any non-human mammal.

[0388] The modified cells or animals are useful in the study of gene function and regulation. For example, a series of small deletions and/or substitutions can be made in the host's native gene to determine the role of different exons in biological processes such as oncogenesis or signal transduction. Of interest is the use of genes to construct transgenic animal models for cancer, where expression of the subject protein is specifically reduced or absent. Specific constructs of interest include anti-sense constructs, which will block expression, expression of dominant negative mutations, and gene over-expression.

[0389] One can also provide for expression of the gene, *e.g.*, a subject gene, or variants thereof, in cells or tissues where it is not normally expressed, at levels not normally present in such cells or tissues, or at abnormal times of development. One can also generate host cells (including host cells in transgenic animals) that comprise a heterologous nucleic

acid molecule which encodes a polypeptide which functions to modulate expression of an endogenous promoter or other transcriptional regulatory region, or the biological activity of a subject polypeptide. The transgenic animals can also be used in functional studies, for example drug screening, to determine the effect of a candidate drug on a biological activity of

Examples

[0390] The examples, which are intended to be purely exemplary of the invention and should therefore not be considered to limit the invention in any way, also describe and detail aspects and embodiments of the invention discussed above. The examples are not intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (*e.g.*, amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

[0391] While the present invention has been described with reference to the specific embodiments thereof, it should be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the true spirit and scope of the invention. In addition, many modifications can be made to adapt a particular situation, material, composition of matter, process, process step or steps, to the objective, spirit and scope of the present invention. All such modifications are intended to be within the scope of the claims appended hereto.

[0392] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed. Moreover, it must be understood that the invention is not limited to the particular embodiments described, as such may, of course, vary. Further, the terminology used to describe particular embodiments is not intended to be limiting, since the scope of the present invention will be limited only by its claims.

[0393] With respect to ranges of values, the invention encompasses each intervening value between the upper and lower limits of the range to at least a tenth of the lower limit's unit, unless the context clearly indicates otherwise. Further, the invention encompasses any other stated intervening values. Moreover, the invention also encompasses ranges excluding either or both of the upper and lower limits of the range, unless specifically excluded from the stated range.

[0394] Unless defined otherwise, the meanings of all technical and scientific terms used herein are those commonly understood by one of ordinary skill in the art to which this invention belongs. One of ordinary skill in the art will also appreciate that any methods and materials similar or equivalent to those described herein can also be used to practice or test the invention. Further, all publications mentioned herein are incorporated by reference.

[0395] It must be noted that, as used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a subject polypeptide" includes a plurality of such polypeptides and reference to "the agent" includes reference to one or more agents and equivalents thereof known to those skilled in the art, and so forth.

[0396] Further, all numbers expressing quantities of ingredients, reaction conditions, % purity, polypeptide and polynucleotide lengths, and so forth, used in the specification and claims, are modified by the term "about," unless otherwise indicated. Accordingly, the numerical parameters set forth in the specification and claims are approximations that may vary depending upon the desired properties of the present invention. At the very least, and not as an attempt to limit the application of the doctrine of equivalents to the scope of the claims, each numerical parameter should at least be construed in light of the number of reported significant digits, applying ordinary rounding techniques. Nonetheless, the numerical values set forth in the specific examples are reported as precisely as possible. Any numerical value, however, inherently contains certain errors from the standard deviation of its experimental measurement.

[0397] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

Example 1 Expression in *E. coli*

[0398] Sequences can be expressed in *E. coli*. Any one or more of the sequences according to SEQ. ID. NOS. 1-187 and 375-484 can be expressed in *E. coli* by subcloning the entire coding region, or a selected portion thereof, into a prokaryotic expression vector. For example, the expression vector pQE16 from the QIA expression prokaryotic protein expression system (Qiagen, Valencia, CA) can be used. The features of this vector that make it useful for protein expression include an efficient promoter (phage T5) to drive

transcription, expression control provided by the lac operator system, which can be induced by addition of IPTG (isopropyl-beta-D-thiogalactopyranoside), and an encoded 6XHis tag coding sequence. The latter is a stretch of six histidine amino acid residues which can bind very tightly to a nickel atom. This vector can be used to express a recombinant protein with a 6XHis. tag fused to its carboxyl terminus, allowing rapid and efficient purification using Ni-coupled affinity columns.

[0399] The entire or the selected partial coding region can be amplified by PCR, then ligated into digested pQE16 vector. The ligation product can be transformed by electroporation into electrocompetent *E. coli* cells (for example, strain M15[pREP4] from Qiagen), and the transformed cells may be plated on ampicillin-containing plates. Colonies may then be screened for the correct insert in the proper orientation using a PCR reaction employing a gene-specific primer and a vector-specific primer. Also, positive clones can be sequenced to ensure correct orientation and sequence. To express the proteins, a colony containing a correct recombinant clone can be inoculated into L-Broth containing 100 µg/ml of ampicillin, and 25 µg/ml of kanamycin, and the culture allowed to grow overnight at 37 degrees C. The saturated culture may then be diluted 20-fold in the same medium and allowed to grow to an optical density of 0.5 at 600 nm. At this point, IPTG can be added to a final concentration of 1 mM to induce protein expression. After growing the culture for an additional 5 hours, the cells may be harvested by centrifugation at 3000 times *g* for 15 minutes.

[0400] The resultant pellet can be lysed with a mild, nonionic detergent in 20 mM Tris HCl (pH 7.5) (B PER.TM. Reagent from Pierce, Rockford, IL), or by sonication until the turbid cell suspension turns translucent. The resulting lysate can be further purified using a nickel-containing column (Ni-NTA spin column from Qiagen) under non-denaturing conditions. Briefly, the lysate will be adjusted to 300 mM NaCl and 10 mM imidazole, then centrifuged at 700 times *g* through the nickel spin column to allow the His-tagged recombinant protein to bind to the column. The column will be washed twice with wash buffer (for example, 50 mM NaH₂ PO₄, pH 8.0; 300 mM NaCl; 20 mM imidazole) and eluted with elution buffer (for example, 50 mM NaH₂ PO₄, pH 8.0; 300 mM NaCl; 250 mM imidazole). All the above procedures will be performed at 4 degrees C. The presence of a purified protein of the predicted size can be confirmed with SDS-PAGE.

Example 2: Expression in Mammalian Cells

[0401] The sequences encoding the proteins of Example 1 can be cloned into the pENTR vector (Invitrogen) by PCR and transferred to the mammalian expression vector pDEST12.2 per manufacturer's instructions (Invitrogen). Introduction of the recombinant construct into the host cell can be effected by transfection with Fugene 6 (Roche) per manufacturer's instructions. The host cells containing one of polynucleotides of the invention can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF). A number of types of cells can act as suitable host cells for expression of the proteins. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Example 3: Expression in Cell-Free Translation Systems

[0402] Cell-free translation systems can also be employed to produce proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors containing SP6 or T7 promoters for use with prokaryotic and eukaryotic hosts have been described (Sambrook et al., 1989). These DNA constructs can be used to produce proteins in a rabbit reticulocyte lysate system or in a wheat germ extract system.

[0403] Specific expression systems of interest include plant, bacterial, yeast, insect cell and mammalian cell derived expression systems. Expression systems in plants include those described in U.S. Patent No. 6,096,546 and U.S. Patent No. 6,127,145. Expression systems in bacteria include those described by Chang et al., 1978, Goeddel et al., 1979, Goeddel et al., 1980, EP 0 036,776, U.S. Patent No. 4,551,433; DeBoer et al., 1983, and Siebenlist et al., 1980.

[0404] Mammalian expression is further accomplished as described in Dijkema et al. 1985, Gorman et al., 1982, Boshart et al., 1985, and U.S. Patent No. 4,399,216. Other features of mammalian expression are facilitated as described in Ham and Wallace, Meth. Enz., 1979, Barnes and Sato, 1980, U.S. Patent Nos. 4,767,704, 4,657,866, 4,927,762, 4,560,655, WO 90/103430, WO 87/00195, and U.S. RE 30,985.

Example 4: Expression

[0405] Primers can be designed to amplify the secreted factors using PCR and cloned into pENTR/D-TOPO vectors (Invitrogen, Carlsbad, CA). The secreted factors in pENTR/D-

TOPO can be cloned into the yeast expression vector pYES-DEST52 by Gateway LR reaction (Invitrogen, Carlsbad, CA). The resulting yeast expression vectors can be transformed into INVSc1 strain from Invitrogen to express the secreted factors according to the manufacturer's protocol (Invitrogen, Carlsbad CA). The expressed secreted factors will have a 6XHis tag at the C-terminal. Expressed protein can be purified with ProBond™ resin (Invitrogen, Carlsbad, CA).

[0406] Expression systems in yeast include those described in Hinnen et al., 1978, Ito et al., 1983, Kurtz et al., 1986, Kunze et al., 1985, Gleeson et al., 1986, Roggenkamp et al., 1986, Das et al., 1984, De Louvencourt et al., 1983, Van den Berg et al., 1990, Kunze et al., 1985, Cregg et al. 1985, U.S. Patent No. 4,837,148, U.S. Patent No. 4,929,555, Beach and Nurse, 1981, Davidow et al., 1985, Gaillardin et al., 1985, Ballance et al., 1983, Tilburn et al., 1983, Yelton et al., 1984, Kelly and Hynes, 1985, EP 0 244,234, and WO 91/00357.

Example 5: Expression in a Baculovirus Expression System.

[0407] The secreted factors in pENTR/D-TOPO can be cloned into Baculovirus expression vector pDEST10 by Gateway LR reaction (Invitrogen, Carlsbad, CA). The secreted factors can be expressed by the Bac-to-Bac expression system from Invitrogen (Carlsbad CA), briefly described as follows. The expression vectors containing the secreted factors are transformed into competent DH10Bac™ *E. coli* strain and selected for transposition. The resulting *E. coli* contain recombinant bacmid that contains the secreted factor. High molecular weight DNA can be isolated from the *E. coli* containing the recombinant bacmid and then transfected into insect cells with Cellfectin reagent. The expressed secreted factors will have a 6XHis tag at N-terminal. Expressed protein will be purified by ProBond™ resin (Invitrogen, Carlsbad, CA).

[0408] Expression of heterologous genes in insects can be accomplished as described in U.S. Patent No. 4,745,051; Doerfler et al., 1987; Friesen *et al.*, 1986; EP 0 127,839, EP 0 155,476, Vlak *et al.*, 1988, Miller *et al.*, 1988, Carbonell *et al.*, 1988, Maeda *et al.*, 1985, Lebacq-Verheyden *et al.*, 1988, Smith *et al.*, 1985, Miyajima *et al.*; and Martin *et al.*, 1988. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts have been previously described (Setlow et al., 1986, Luckow *et al.*, 1988; Miller *et al.*, 1986; Maeda *et al.*, 1985).

Example 6: Primer Design

[0409] To design the forward primer for PCR amplification, the melting point of the first 20 to 24 bases of the primer can be calculated by counting total A and T residues, then

multiplying by 2. To design the reverse primer for PCR amplification, the melting point of the first 20 to 24 bases of the reverse complement, with the sequences written from 5-prime to 3-prime can be calculated by counting the total G and C residues, then multiplying by 4. Both start and stop codons can be present in the final amplified clone. The length of the primers is such to obtain melting temperatures within 63 degrees C to 68 degrees C. Adding the bases "CACC" to the forward primer renders it compatible for cloning the PCR product with the TOPO pENTR/D (Invitrogen, CA).

Example 7: Reverse Transcriptase Reaction

[0410] cDNA can be prepared by the following method. Between 200 ng and 1.0 µg mRNA is added to 2 µl DMSO and the volume adjusted to 11 µl with DEPC-treated water. One µl Oligo dT is added to the tube, and the mixture is heated at 70° C for 5 min., quickly chilled on ice for 2 min., and the mixture is collected at the bottom of the tube by brief centrifugation. The following 1st strand components are then added to the mRNA mixture: 2 µl 10X Stratascript (Stratagene, CA) 1st strand buffer, 1 µl 0.1 M DTT, 1 µl 10 mM dNTP mix (10 mM each of dG, dA, dT and dCTP), 1 µl RNase inhibitor, 3 µl Stratascript RT (50 U/ µl). The contents are gently mixed and the mixture collected by brief centrifugation. The mixture is incubated in a 42° C water bath for 1 hour, placed in a 70° C water bath for 15 min. to stop the reaction, transferred to ice for 2 min., and centrifuged briefly in a microfuge to collect the reaction product at the bottom of the reaction vessel. Two µl RNase H is then added to the tube, the contents are mixed well, incubated at 37° C in a water bath for 20 min., and centrifuged briefly in a microfuge to collect the reaction product at the bottom of the reaction vessel. The reaction mixture can proceed directly to PCR or be stored at – 20° C.

Example 8: Full Length PCR

[0411] Full length PCR can be achieved by placing the products of the reaction described in Example 7, with primers diluted to 5µM in water, into a reaction vessel and adding a reaction mixture composed of 1x Taq buffer, 25 mM dNTP, 10 ng cDNA pool, TaqPlus (Stratagene, CA) (5u/ul), PfuTurbo (Stratagene, CA) (2.5u/ul), water. The contents of the reaction vessel are then mixed gently by inversion 5-6 times, placed into a reservoir where 2µl F₁/R₁ primers are added, the plate sealed and placed in the thermocycler. The PCR reaction is comprised of the following eight steps. Step 1: 95° C for 3 min. Step 2: 94° C for 45 sec. Step 3: 0.5° C/sec to 56-60° C. Step 4: 56-60° C for 50 sec. Step 5: 72° C for 5 min. Step 6: Go to step 2, perform 35-40 cycles. Step 7: 72° C for 20 min. Step 8: 4° C.

[0412] The products can then be separated on a standard 0.8 to 1.0% agarose gel at 40 to 80 V, the bands of interest excised by cutting from the gel, and stored at -20° C until extraction. The material in the bands of interest can be purified with QIAquick 96 PCR Purification Kit (Qiagen, CA) according to the manufacturer instructions. Cloning can be performed with the Topo Vector pENTR/D-TOPO vector (Invitrogen, CA) according to the manufacturer's instructions.

References

[0413] The specification is most thoroughly understood in light of the following references, all of which are hereby incorporated by reference in their entireties. The disclosures of the patents and other references cited above are also hereby incorporated by reference.

1. Agou, F., Quevillon, S., Kerjan, P., Latreille, M.T., Mirande, M. (1996) Functional replacement of hamster lysyl-tRNA synthetase by the yeast enzyme requires cognate amino acid sequences for proper tRNA recognition. *Biochemistry* 35:15322-15331.
2. Agrawal, S., Crooke, S.T. eds. (1998) Antisense Research and Application (Handbook of Experimental Pharmacology, Vol 131). Springer-Verlag New York, Inc.
3. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D. (1994) Molecular Biology of the Cell. 3rd ed. Garland Publishing, Inc.
4. Alexander, D.R. (2000) The CD45 tyrosine phosphatase: a positive and negative regulator of immune cell function. *Semin. Immunol* 12:349-359.
5. Allison, A.C. (2000) Immunosuppressive drugs: the first 50 years and a glance forward. *Immunopharmacology* 47:63-83.
6. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic alignment search tool. *J. Mol. Biol.* 215:403-410.
7. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zheng, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
8. Amor, J.C., Harrison, D.H., Kahn, R.A., Ringe, D. (1994) Structure of the human ADP-ribosylation factor 1 complexed with GDP. *Nature* 372:704-708.
9. Andreeff, M., Pinkel, D. eds. (1999) Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications. John Wiley & Sons.

10. Andres, D.A., Shao, H., Crick, D.C., Finlin, B.S. (1997) Expression cloning of a novel farnesylated protein, RDJ2, encoding a DnaJ protein homologue. *Arch. Biochem. Biophys.* 346:113-124.
11. Ansel, H.C., Allen, L., Popovich, N.G. eds. (1999) Pharmaceutical Dosage Forms and Drug Delivery Systems. 7th ed. Lippencott Williams and Wilkins Publishers.
12. Attardi, L.D., Jacks, T. (1999) The role of p53 in tumor suppression: lessons from mouse models. *Cell Mol. Life Sci.* 55:48-63.
13. Aubry, M., Marineau, C., Zhang, F.R., Zahed, L., Figlewicz, D., Delattre, O., Thomas, G., de Jong, P.J., Julien, J.P., Rouleau, G.A. (1992) Cloning of six new genes with zinc finger motifs mapping to short and long arms of human acrocentric chromosome 22 (p and q11.2). *Genomics* 13:641-648.
14. Ausubel, F., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., eds. (1999) Short Protocols in Molecular Biology. 4th ed. Wiley & Sons.
15. Baksh, S., Burakoff, S.J. (2000) The role of calcineurin in lymphocyte activation. *Semin. Immunol.* 12:405-415.
16. Ballance, D.J., Buxton, F.P., Turner, G. (1983) Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*. *Biochem. Biophys. Res. Commun.* 112:284-289.
17. Barany, F. (1985) Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering. *Gene* 37:111-123.
18. Barnes, D., Sato, G. (1980) Methods for growth of cultured cells in serum-free medium. *Anal. Biochem.* 102:255-270.
19. Barton, M.C., Hoekstra, M.F., Emerson, B.M. (1990) Site-directed, recombination-mediated mutagenesis of a complex gene locus. *Nucleic Acids Res.* 18:7349-7355.
20. Bashkin, J.K., Sampath, U., Frolova, E. (1995) Ribozyme mimics as catalytic antisense reagents. *Appl. Biochem. Biotechnol.* 54:43-56.
21. Bassett, D.E., Eisen, M.B., Boguski, M.S. (1999) Gene expression informatics - it's all in your mine. *Nature Genetics* 21:51-55.
22. Bast, R.C., Kufe, D.W., Pollock, R.E., Weichselbaum, R.R., Holland, J.F., Frei, E., eds. (2000) Cancer Medicine. 5th ed. B.C. Decker, Inc.
23. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Eddy, S.R., Griffiths-Jones, S., Howe, K.L., Marshall, M., Sonnhammer, E.L.L. (2000) The Pfam protein families database. *Nucleic Acids Research* 30:276-280.

24. Battini, R., Ferrari, S., Kaczmarek, L., Calabretta, B., Chen, S.T., Baserga, R. (1987) Molecular cloning of a cDNA for a human ADP/ATP carrier which is growth-regulated. *J. Biol. Chem.* 262:4355-4359.
25. Beach, D., Durkacz, B., Nurse, P. (1982) Functionally homologous cell cycle control genes in budding and fission yeast. *Nature* 300:706-709.
26. Beigelman, L., Karpeisky, A., Matulic-Adamic, J., Haeberli, P., Sweedler, D., Usman, N. (1995) Synthesis of 2'-modified nucleotides and their incorporation into hammerhead ribozymes. *Nucleic Acids Res.* 23:4434-4442.
27. Bennett, J. (2000) Gene therapy for retinitis pigmentosa. *Curr. Opin. Mol. Ther.* 2:420-425.
28. Berinstein, N.L. (2002) Carcinoembryonic antigen as a target for therapeutic anticancer vaccines: a review. *J. Clin. Oncol.* 20:2197-2207.
29. Bibikova, M., Beumer, K., Trautman, J.K., Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* 300:764.
30. Birney, E., Durbin, R. (2000) Using GeneWise in the *Drosophila* annotation experiment. *Genome Res.* 10:547-548.
31. Blackwell, J.M., Barton, C.H., White, J.K., Searle, S., Baker, A.M., Williams, H., Shaw, M.A. (1995) Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. *Mol. Med.* 1:194-205.
32. Bodzioch, M., Orso, E., Klucken, J., Langmann, T., Bottcher, A., Diederich, W., Drobnik, W., Barlage, S., Buchler, C., Porsch-Ozcurumez, M., Kaminski, W.E., Hahmann, H.W., Oette, K., Rothe, G., Aslanidis, C., Lackner, K.J., Schmitz, G. (1999) The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nat. Genet.* 1999 22:347-351.
33. Bonifaci, N., Moroianu, J., Radu, A., Blobel, G. (1997) Karyopherin beta2 mediates nuclear import of a mRNA binding protein. *Proc. Natl. Acad. Sci.* 94:5055-5060.
34. Bono, H., Kasukawa, T., Furuno, M., Hayashizaki, Y., Okazaki, Y. (2002) FANTOM DB: database of Functional Annotation of RIKEN Mouse cDNA Clones. *Nucleic Acids Res.* 30:116-118.
35. Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B., Schaffner, W. (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* 41:521-530.

36. Bowtell, D.D.L. (1999) Options available - from start to finish - for obtaining expression data by microarray. *Nature Genetics* 21:25-32.
37. Brenner, S., Williams, S.R., Vermass, E.H., Storck, T., Moon, K., McCollum, C., Mao, J.I., Luo, S., Kirchner, J.J., Eletr, S., DuBridge, R.B., Burcham, T., Albrecht, G. (2000) *In vitro* cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* 97:1665-1670.
38. Brinster, R.L., Chen, H.Y., Trumbauer, M., Senear, A.W., Warren, R., Palmiter, R.D. (1981) Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27:223-231.
39. Brock, G. (2000) Sildenafil citrate (Viagra®). *Drugs Today* 36:125-134.
40. Brown, J.R., Daar, I.O., Krug, J.R., Maquat, L.E. (1985) Characterization of the functional gene and several processed pseudogenes in the human triosephosphate isomerase gene family. *Mol. Cell Biol.* 5:1694-1706.
41. Brown, P.O., Botstein, D. (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genetics* 21:33-37.
42. Brunelleschi, S., Penengo, L., Santoro, M.M., Gaudino, G. (2002) Receptor tyrosine kinases as target for anti-cancer therapy. *Curr. Pharm. Des.* 8:1959-1972.
43. Brutlag, D.L., Dautricourt, J.P., Diaz, R., Fier, J., Moxon, B., Stamm, R. (1993). BLAZE: An implementation of the Smith-Waterman comparison algorithm on a massively parallel computer. *Computers and Chemistry* 17:203-207.
44. Capecchi, M.R. (1989) The new mouse genetics: altering the genome by gene targeting. *Trends Genet.* 5:70-76. Sasaki, T., Fujimoto, K., Sakai, K., Nemoto, M., Nakai, N., Tajima, N. (2003) Gene and cell-based therapy for diabetes mellitus: endocrine gene therapeutics. *Endocr. Pathol.* 14:141-144.
45. Carbonell, L.F., Hodge, M.R., Tomalski, M.D., Miller, L.K. (1988) Synthesis of a gene coding for an insect-specific scorpion neurotoxin and attempts to express it using baculovirus vectors. *Gene* 73:409-418.
46. Chakravarty, A. (1999) Population genetics - making sense out of sequence. *Nature Genetics* 21:56-60.
47. Chalifour, L.E., Fahmy, R., Holder, E.L., Hutchinson, E.W., Osterland, C.K., Schipper, H.M., Wang, E. (1994) A method for analysis of gene expression patterns. *Anal. Biochem.* 216: 299-304.

48. Chalut, C., Gallois, Y., Poterszman, A., Moncollin, V., Egly, J.M. (1995) Genomic structure of the human TATA-box-binding protein (TBP). *Gene* 161:277-282.
49. Chang, A.C., Nunberg, J.H., Kaufman, R.J., Erlich, H.A., Schimke, R.T., Cohen, S.N. (1978) Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dihydrofolate reductase. *Nature* 275:617-624.
50. Chang, M.S., Chang, C.L., Huang, C.J., Yang, Y.C. (2000) p29, a novel GCIP-interacting protein, localizes in the nucleus. *Biochem. Biophys. Res. Commun.* 279:732-737.
51. Chen, F.W., Ioannou, Y.A. (1998) Ribosomal proteins in cell proliferation and apoptosis. *Int. Rev. Immunol.* 18:429-448.
52. Chen, S.Y., Bagley, J., Marasco, W.A. (1994) Intracellular antibodies as a new class of therapeutic molecule for gene therapy. *Hum. Gene Ther.* 5:595-601.
53. Cheng, W.F., Hung, C.F., Chai, C.Y., Hsu, K.F., He, L., Ling, M., Wu, T.C. (2001) Tumor-specific immunity and angiogenesis generated by a DNA vaccine encoding calreticulin linked to a tumor antigen. *J. Clin. Invest.* 108:669-678.
54. Cheung, V.G., Morley, M., Aquilar, F., Massimi, A., Kucherlapati, R., Childs, G. (1999) Making and reading microarrays. *Nature Genetics* 21:15-19.
55. Chien, C., Bartel, P.L., Sternglanz, R., Fields S. (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci.* 88:9578-9581.
56. Christa, L., Simon, M.T., Flinois, J.P., Gebhardt, R., Brechot, C., Lasserre, C. (1994) Overexpression of glutamine synthetase in human primary liver cancer. *Gastroenterology* 106:1312-1320.
57. Clark, C.M., Karlawish, J.H. (2003) Alzheimer disease: current concepts and emerging diagnostic and therapeutic strategies. *Ann. Intern. Med.* 138:400-410.
58. Clark, H.F., Gurney, A.L., Abaya, E., Baker, K., Baldwin, D., Brush, J., Chen, J., Chow, B., Chui, C., Crowley, C., Currell, B., Deuel, B., Dowd, P., Eaton, D., Foster, J., Grimaldi, C., Gu, Q., Hass, P.E., Heldens, S., Huang, A., Kim, H.S., Klimowski, L., Jin, Y., Johnson, S., Lee, J., Lewis, L., Liao, D., Mark, M., Robbie, E., Sanchez, C., Schoenfeld, J., Seshagiri, S., Simmons, L., Singh, J., Smith, V., Stinson, J., Vagts, A., Vandlen, R., Watanabe, C., Wieand, D., Woods, K., Xie, M.H., Yansura, D., Yi, S., Yu, G., Yuan, J., Zhang, M., Zhang, Z., Goddard, A., Wood, W.I., Godowski, P. (2003) The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to

- Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment. *Genome Res.* 13:2265-2270; Epublised Sep 15, 2003.
59. Cole, K.A., Krizman, D.B., Emmert-Buck, M.R. (1999) The genetics of cancer - a 3D model. *Nature Genetics* 21:38-41.
 60. Colicelli, J., Lobel, L.I., Goff, S.P. (1985) A temperature-sensitive mutation constructed by "linker insertion" mutagenesis. *Mol. Gen. Genet.* 199:537-539.
 61. Collins, F.S. (1999) Microarrays and macroconsequences. *Nature Genetics* 21:2.
 62. Comuzzie, A.G., Allison, D.B. (1998) The search for human obesity genes. *Science* 280:1374-1377.
 63. Cormand, B., Montfort, M., Chabas, A., Vilageliu, L., Grinberg, D. (1997) Genetic fine localization of the beta-glucocerebrosidase (GBA) and prosaposin (PSAP) genes: implications for Gaucher disease. *Hum. Genet.* 100:75-79.
 64. Cregg, J.M., Barringer, K.J., Hessler, A.Y., Madden, K.R. (1985) *Pichia pastoris* as a host system for transformations. *Mol. Cell. Biol.* 5:3376-3385.
 65. Crooke, S.T. (1996) Progress in antisense therapeutics. *Med. Res. Rev.* 16:319-344.
 66. Crouch, R.J. (1990) Ribonuclease H: from discovery to 3D structure. *New Biol.* 2:771-777.
 67. Curcio, L.D., Bouffard, D.Y., Scanlon, K.J. (1997) Oligonucleotides as modulators of cancer gene expression. *Pharmacol. Ther.* 74:317-332.
 68. Das, S., Kellermann, E., Hollenberg, C.P. (1984) Transformation of *Kluyveromyces fragilis*. *J. Bacteriol.* 158:1165-1167.
 69. Davidow, L.S., Kaczmarek, F.S., DeZeeuw, J.R., Conlon, S.W., Lauth, M.R., Pereira, D.A., Franke, A.E. (1987) The *Yarrowia lipolytica* LEU2 gene. *Curr. Genet.* 11:377-383.
 70. de Boer, H.A., Comstock, L.J., Vasser, M. (1993) The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc. Natl. Acad. Sci.* 80:21-25.
 71. De Louvencourt, L., Fukuhara, H., Heslot, H., Wesolowski, M. (1983) Transformation of *Kluyveromyces lactis* by killer plasmid DNA. *J. Bacteriol.* 154:737-742.
 72. Deasy, B.M., Huard, J. (2002) Gene therapy and tissue engineering based on muscle-derived stem cells. *Curr. Opin. Mol. Ther.* 4:382-389.

73. Delahunty, C., Ankener, W., Deng, Q., Eng, J., Nickerson, D.A. (1996) Testing the feasibility of DNA typing for human identification by PCR and an oligonucleotide ligation assay. *Am. J. Human Genetics* 58:1239-1246.
74. Deutscher, M.P., Simon, M.I., Abelson, J.N., eds. (1990) Guide to Protein Purification: Methods in Enzymology. (Methods in Enzymology Series, Vol 182). Academic Press.
75. Dieffenbach, C.W., Dveksler, G.S., eds. (1995) PCR Primer: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
76. Dijkema, R., van der Meide, P.H., Pouwels, P.H., Caspers, M., Dubbeld, M., Schellekens, H. (1985) Cloning and expression of the chromosomal immune interferon gene of the rat. *EMBO J.* 4:761-767.
77. Doerfler, W., Bohm, P., eds. (1987) The Molecular Biology Of Baculoviruses. Springer-Verlag, Inc.
78. Doll, A., Grzeschik, K.H. (2001) Characterization of two novel genes, WBSCR20 and WBSCR22, deleted in Williams-Beuren syndrome. *Cytogenet. Cell Genet.* 95:20-27.
79. Doolittle, R.F., Abelson, J.N., Simon, M.I., eds. (1996) Computer Methods for Macromolecular Sequence Analysis. 1st ed. Academic Press.
80. Ducrest, A.L., Suzutorisz, H., Lingner, J., Nabholz, M. (2002) Regulation of the human telomerase reverse transcriptase gene. *Oncogene* 21:541-52.
81. Dutoit, V., Taub, R.N., Papadopoulos, K.P., Talbot, S., Keohan, M.L., Brehm, M., Gnjjatic, S., Harris, P.E., Bisikirska, B., Guillaume, P., Cerottini, J.C., Hesdorffer, C.S., Old, L.J., Valmori, D. (2002) Multiepitope CD8⁺ T cell response to an NY-ESO-1 peptide vaccine results in imprecise tumor targeting. *J. Clin. Invest.* 110:1813-1822.
82. Egilsson, V., Gudnason, V., Jonasdottir, A., Ingvarsson, S., Andresdottir, V. (1986) Catabolite repressive effects of 5-thio-D-glucose on *Saccharomyces cerevisiae*. *J. Gen. Microbiol.* 132:3309-3313.
83. Ehrhardt, G.R., Korherr, C., Wieler, J.S., Knaus, M., Schrader, J.W. (2001) A novel potential effector of M-Ras and p21 Ras negatively regulates p21 Ras-mediated gene induction and cell growth. *Oncogene* 20:188-197.

84. Espejo, A., Cote, J., Bednarek, A., Richard, S., Bedford, M.T. (2002) A protein-domain microarray identifies novel protein-protein interactions. *Biochem. J.* 367:697-702.
85. Everett, R.D., Meredith, M., Orr, A., Cross, A., Kathoria, M., Parkinson, J. (1997) A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.* 16:1519-1530.
86. Fanning, A.S., Anderson, J.M. (1999) Protein modules as organizers of membrane structure. *Curr. Opin. Cell Biol.* 11:432-439.
87. Fields, S., Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340:245-246.
88. Fiers, W. (1991) Tumor necrosis factor. Characterization at the molecular, cellular and *in vivo* level. *FEBS Lett.* 285:199-212.
89. Fisch, P., Forster, A., Sherrington, P.D., Dyer, M.J., Rabbitts, T.H. (1993) The chromosomal translocation t(X;14)(q28;q11) in T-cell pro-lymphocytic leukaemia breaks within one gene and activates another. *Oncogene* 8:3271-3276.
90. Fishman, P.S., Oyler, G.A. (2002) Significance of the parkin gene and protein in understanding Parkinson's disease. *Curr. Neurol. Neurosci. Rep.* 2:296-302.
91. Forgac, M. (1999) Structure and properties of the vacuolar (H⁺)-ATPases. *J. Biol. Chem.* 274:12,951-12,954.
92. Frank, I. (2002) Antivirals against HIV-1. *Clin. Lab. Med.* 22:741-757.
93. Frithz, G., Ericsson, P., Ronquist, G. (1976) Serum adenylate kinase activity in the early phase of acute myocardial infarction. *Ups J Med Sci.* 81:155-158.
94. Funakoshi, I., Kato, H., Horie, K., Yano, T., Hori, Y., Kobayashi, H., Inoue, T., Suzuki, H., Fukui, S., Tsukahara, M., et al. (1992) Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase. *Arch. Biochem. Biophys.* 295:180-187.
95. Furth, P.A., Shamay, A., Wall, R.J., Hennighausen, L. (1992) Gene transfer into somatic tissues by jet injection. *Anal. Biochem.* 205:365-368.
96. Gaillardin, C., Ribet, A.M. (1987) LEU2 directed expression of beta-galactosidase activity and phleomycin resistance in *Yarrowia lipolytica*. *Curr. Genet.* 11:369-375.
97. Gao, X., Nawaz, Z. (2002) Progesterone receptors - animal models and cell signaling in breast cancer: Role of steroid receptor coactivators and corepressors of progesterone receptors in breast cancer. *Breast Cancer Res.* 4:182-186.

98. Gao, Y., Melki, R., Walden, P.D., Lewis, S.A., Ampe, C., Rommelaere, H., Vandekerckhove, J., Cowan, N.J. (1994) A novel cochaperonin that modulates the ATPase activity of cytoplasmic chaperonin. *J. Cell Biol.* 125:989-996.
99. Gaudilliere, B., Shi, Y., Bonni, A. (2002) RNA interference reveals a requirement for MEF2A in activity-dependent neuronal survival. *J. Biol. Chem.* 277:46,442-46,446.
100. Gavrieli, Y., Sherman, Y., Ben-Sasson, S.A. (1992) Identification of programmed cell death *in situ* via specific labeling of nuclear DNA fragmentation. *J. Cell Biol.* 119:493-501.
101. Geffen D.B., Man S. (2002) New drugs for the treatment of cancer, 1990-2001. *Isr. Med. Assoc. J.* 4:1124-31.
102. Gennaro, A., ed. (2000) Remington: The Science and Practice of Pharmacy. 20th ed. Lippincott, Williams, & Wilkins.
103. Ghofrani, H.A., Rose, F., Schermuly, R.T., Olschewski, H., Wiedemann, R., Kreckel, A., Weissmann, N., Ghofrani, S., Enke, B., Seeger, W., Grimminger, F. (2003) Oral sildenafil as long-term adjunct therapy to inhaled iloprost in severe pulmonary arterial hypertension. *J. Am. Coll. Cardiol.* 42:158-164.
104. Gillingham, A.K., Pfeifer, A.C., Munro, S. (2002) CASP, the alternatively spliced product of the gene encoding the CCAAT-displacement protein transcription factor, is a Golgi membrane protein related to giantin. *Mol. Biol. Cell* 13:3761-3774.
105. Gingras, M.C., Lapillonne, H., Margolin, J.F. (2002) TREM-1, MDL-1, and DAP12 expression is associated with a mature stage of myeloid development. *Mol. Immunol.* 38:817-824.
106. Girschick, H.J., Grammer, A.C., Nanki, T., Vazquez, E., Lipsky, P.E. (2002) Expression of recombination activating genes 1 and 2 in peripheral B cells of patients with systemic lupus erythematosus. *Arthritis. Rheum.* 46:1255-1263.
107. Gmeiner, W.H., Horita, D.A. (2001) Implications of SH3 domain structure and dynamics for protein regulation and drug design. *Cell Biochem. Biophys.* 35:127-140.
108. Goeddel, D.V., Heyneker, H.L., Hozumi, T., Arentzen, R., Itakura, K., Yansura, D.G., Ross, M.J., Mizzari, G., Crea, R., Seeburg, P.H. (1979) Direct expression in *E. coli* of a DNA sequence coding for human growth hormone. *Nature* 281:544-548.

109. Goldstein, L.S.B., Yang, Z. (2000) Microtubule-based transport systems in neurons: the roles of kinesins and dyneins. *Annu. Rev. Neurosci.* 23:39-71.
110. Golovkina, T.V., Chervonsky, A., Dudley, J.P., Ross, S.R. (1992) Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell* 69:637-645.
111. Gonnet, G.H., Cohen, M.A., Benner, S.A. (1992) Exhaustive matching of the entire protein sequence database. *Science* 256:1443-1445.
112. Gordan, J.D., Vonderheide, R.H. (2002) Universal tumor antigens as targets for immunotherapy. *Cytotherapy* 4:317-327.
113. Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I., Howard, B.H. (1982) The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eucaryotic cells by DNA-mediated transfection. *Proc. Natl. Acad. Sci.* 79:6777-6781.
114. Gray, T.A., Hernandez, L., Carey, A.H., Schaldach, M.A., Smithwick, M.J., Rus, K.M., Graves, J.A., Stewart, C.L., Nicholls, R.D. (2002) The ancient source of a distinct gene family encoding proteins featuring RING and C(3)H zinc-finger motifs with abundant expression in developing brain and nervous system. *Genomics.* 66:76-86.
115. Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M. (1999) Introduction to Genetic Analysis. 7th ed. W.H. Freeman.
116. Griffiths, M., Beaumont, N., Yao, S.Y., Sundaram, M., Boumah, C.E., Davies, A., Kwong, F.Y., Coe, I., Cass, C.E., Young, J.D., Baldwin, S.A. (1997) Cloning of a human nucleoside transporter implicated in the cellular uptake of adenosine and chemotherapeutic drugs. *Nat. Med.* 3:89-93.
117. Grosschedl, R., Baltimore, D. (1985) Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. *Cell* 41:885-897.
118. Grosveld, F., Kollias, G., eds. (1992) Transgenic Animals. 1st ed. Academic Press.
119. Gustin, K., Burk, R.D. (1993) A rapid method for generating linker scanning mutants utilizing PCR. *Biotechniques* 14:22-24.
120. Hacia, J.G. (1999) Resequencing and mutational analysis using oligonucleotide microarrays. *Nature Genetics* 21:42-47.

121. Hadano, S., Yanagisawa, Y., Skaug, J., Fichter, K., Nasir, J., Martindale, D., Koop, B.F., Scherer, S.W., Nicholson, D.W., Rouleau, G.A., Ikeda, J., Hayden, M.R. (2001) Cloning and characterization of three novel genes, ALS2CR1, ALS2CR2, and ALS2CR3, in the juvenile amyotrophic lateral sclerosis (ALS2) critical region at chromosome 2q33-q34: candidate genes for ALS2. *Genomics* 71:200-213.
122. Hall, M., Mickey, D.D., Wenger, A.S., Silverman, L.M. (1985) Adenylate kinase: an oncodevelopmental marker in an animal model for human prostatic cancer. *Clin. Chem.* 31:1689-1691.
123. Ham, R.G., McKeehan, W.L. (1979) Media and growth requirements. *Methods Enzymol.* 58:44-93.
124. Hanada, T., Lin, L., Tibaldi, E.V., Reinherz, E.L., Chishti, A.H. (2000) GAKIN, a novel kinesin-like protein associates with the human homologue of the Drosophila discs large tumor suppressor in T lymphocytes. *J. Biol. Chem.* 275:28,774-28,784.
125. Harlow, E., Lane, D., eds. (1988) Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory.
126. Harlow, E., Lane, D., Harlow, E., eds. (1998) Using Antibodies: A Laboratory Manual: Portable Protocol NO. I. Cold Spring Harbor Laboratory.
127. Hartmann, G., Endres, S., eds. (1999) Manual of Antisense Methodology (Perspectives in Antisense Science). 1st ed. Kluwer Law International.
128. Hassanzadeh, G.H.G., De Silva, K.S., Dambly-Chudiere, C., Brys, L., Ghysen, A., Hamers, R., Muyldermans, S., De Baetselier, P. (1998) Isolation and characterization of single-chain Fv genes encoding antibodies specific for Drosophila Poxn protein. *FEBS Lett.* 437:75-80.
129. Hawes, J.W., Jaskiewicz, J., Shimomura, Y., Huang, B., Bunting, J., Harper, E.T., Harris, R.A. (1996) Primary structure and tissue-specific expression of human beta-hydroxyisobutyryl-coenzyme A hydrolase. *J. Biol. Chem.* 271:26,430-26,434.
130. Heath, J.K., White, S.J., Johnstone, C.N., Catimel, B., Simpson, R.J., Moritz, R.L., Tu, G.F., Ji, H., Whitehead, R.H., Groenen, L.C., Scott, A.M., Ritter, G., Cohen, L., Welt, S., Old, L.J., Nice, E.C., Burgess, A.W. (1997) The human A33 antigen is a transmembrane glycoprotein and a novel member of the immunoglobulin superfamily. *Proc. Natl. Acad. Sci.* 94:469-474.

131. Heiser, A., Coleman, D., Dannull, J., Yancey, D., Maurice, M.A., Lallas, C.D., Dahm, P., Niedzwiecki, D., Gilboa, E., Vieweg, J. (2002) Autologous dendritic cells transfected with prostate-specific antigen RNA stimulate CTL responses against metastatic prostate tumors. *J. Clin. Invest.* 109:409-417.
132. Henningson, C.T. Jr., Stanislaus, M.A., Gewirtz, A.M. (2003) Embryonic and adult stem cell therapy. *J. Allergy Clin. Immunol.* 111:S745-S753.
133. Hinnen, A., Hicks, J.B., Fink, G.R. (1978) Transformation of yeast. *Proc. Natl. Acad. Sci.* 75:1929-1933.
134. Hirsch, D.S., Pirone, D.M., Burbelo, P.D. (2001) A new family of Cdc42 effector proteins, CEPs, function in fibroblast and epithelial cell shape changes. *J. Biol. Chem.* 276:875-883.
135. Ho, L.W., Carmichael, J., Swartz, J., Wyttenbach, A., Rankin, J., Rubinsztein, D.C. (2001) The molecular biology of Huntington's disease. *Psychol. Med.* 31:3-14.
136. Hollis, G.F., Evans, R.J., Stafford-Hollis, J.M., Korsmeyer, S.J., McKearn, J.P. (1989) Immunoglobulin lambda light-chain-related genes 14.1 and 16.1 are expressed in pre-B cells and may encode the human immunoglobulin omega light-chain protein. *Proc. Natl. Acad. Sci.* 86:5552-5556.
137. Homo-Delarche, F. (2001) Is pancreas development abnormal in the non-obese diabetic mouse, a spontaneous model of type I diabetes? *Braz. J. Med. Biol. Res.* 34:437-447.
138. Hong, G.F. (1982) Sequencing of large double-stranded DNA using the dideoxy sequencing technique. *Biosci. Rep.* 2:907-912.
139. Hoogenboom, H.R., de Bruin, A.P., Hufton, S.E., Hoet, R.M., Arends, J.W., Roovers, R.C. (1998) Antibody phage display technology and its applications. *Immunotechnology* 4:1-20.
140. Hooper, M.L. (1993) Embryonal Stem Cells: Introducing Planned Changes into the Animal Germline. Gordon & Breach Science Pub.
141. Hoozemans, J.J., Veerhuis, R., Rozemuller, A.J., Eikelenboom, P. (2002) The pathological cascade of Alzheimer's disease: the role of inflammation and its therapeutic implications. *Drugs Today (Barc)* 38:429-443.
142. Houseman, B.T., Huh, J.H., Kron, S.J., Mrksich, M. (2002) Peptide chips for the quantitative evaluation of protein kinase activity. *Nature Biotechnol.* 20:270-274.

143. Howard, G.C., Bethell, D.R. (2000) Basic Methods in Antibody Production and Characterization. CRC Press.
144. Huynh, D.P., Yang, H.T., Vakharia, H., Nguyen, D., Pulst, S.M. (2003) Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death. *Hum. Mol. Genet.* 12:1485-1496.
145. Ikeda, A., Nishina, P.M., Naggert, J.K. (2002) The tubby-like proteins, a family with roles in neuronal development and function. *J. Cell Sci.* 115(Pt 1):9-14.
146. Ito, H., Fukuda, Y., Murata, K., Kimura, A. (1978) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153:163-168.
147. Jameson, D.M., Sawyer, W.H. (1995) Fluorescence anisotropy applied to biomolecular interactions. *Methods Enzymol.* 246:283-300.
148. Janeway, C.A., Travers, P. Walport, M. Shlomchik, M. (2001) Immunobiology. 5th ed. Garland Publishing.
149. Jeffery, P., Zhu, J. (2002) Mucin-producing elements and inflammatory cells. *Novartis Found. Symp.* 248:51-75, 277-82.
150. Jimbo, T., Kawasaki, Y., Koyama, R., Sato, R., Takada, S., Haraguchi, K., Akiyama, T. (2002) Identification of a link between the tumour suppressor APC and the kinesin superfamily. *Nat. Cell Biol.* 4:323-327.
151. Joberty, G., Perlungher, R.R., Macara, I.G. (1999) The Borgs, a new family of Cdc42 and TC10 GTPase-interacting proteins. *Mol. Cell Biol.* 19:6585-6597.
152. Johns, T.G., Bernard, C.C. (1997) Binding of complement component Clq to myelin oligodendrocyte glycoprotein: a novel mechanism for regulating CNS inflammation. *Mol. Immunol.* 34:33-38.
153. Jolliffe, C.N., Harvey, K.F., Haines, B.P., Parasivam, G., Kumar, S. (2000) Identification of multiple proteins expressed in murine embryos as binding partners for the WW_domains of the ubiquitin-protein ligase Nedd4. *Biochem. J.* 351:557-565.
154. Jones, D.H., Winistorfer, S.C. (1992) Recombinant circle PCR and recombination PCR for site-specific mutagenesis without PCR product purification. *Biotechniques* 12:528-530.
155. Jones, P., ed. (1998a) Vectors: Cloning Applications: Essential Techniques, John Wiley & Son, Ltd.
156. Jones, P., ed. (1998b) Vectors: Expression Systems: Essential Techniques, John Wiley & Son, Ltd.

157. Jorgensen, C., Djouad, F., Apparailly, F., Noel, D. (2003) Engineering mesenchymal stem cells for immunotherapy. *Gene Ther.* 10:928-931.
158. Jost, C.R., Kurucz I., Jacobus, C.M., Titus, J.A., George, A.J., Segal, D.M. (1994) Mammalian expression and secretion of functional single-chain Fv molecules. *J. Biol. Chem.* 269:26,267-26,273.
159. Joulin, V., Richard-Foy, H. (1995) A new approach to isolate genomic control regions. Application to the GATA transcription factor family. *Eur. J. Biochem.* 232:620-626.
160. Jurcic, J.G., Cathcart, K., Pinilla-Ibarz, J., Scheinberg, D.A. (2000) Advances in immunotherapy of hematologic malignancies: cellular and humoral approaches. *Curr. Opin. Hematol.* 7:247-254.
161. Jury, J.A., Perry, A.C., Hall, L. (1999) Identification, sequence analysis and expression of transcripts encoding a putative metalloproteinase, eMDC II, in human and macaque epididymis. *Mol. Hum. Reprod.* 5:1127-1134.
162. Kabat, E.A., Wu T.T. (1991) Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *J. Immunol.* 147:1709-1719.
163. Kamitani, T., Nguyen, H.P., Yeh, E.T. (1997) Preferential modification of nuclear proteins by a novel ubiquitin-like molecule. *J. Biol. Chem.* 272:14,001-14,004.
164. Kantoff, P.W., Halabi, S., Farmer, D.A., Hayes, D.F., Vogelzang, N.A., Small, E.J. (2001) Prognostic significance of reverse transcriptase polymerase chain reaction for prostate-specific antigen in men with hormone-refractory prostate cancer. *J. Clin. Oncol.* 9:3025-3028.
165. Kao, P.N., Chen, L., Brock, G., Ng, J., Kenny, J., Smith, A.J., Cortesy, B. (1994) Cloning and expression of cyclosporin A- and FK506-sensitive nuclear factor of activated T-cells: NF45 and NF90. *J. Biol. Chem.* 269:20,691-20,699.
166. Karanazashvili, G., Abrahamsson, P. (2003) Prostate specific antigen and human glandular kallikrein 2 in early detection of prostate cancer. *J. Urol.* 169:445-457.

167. Kari, C., Chan, T.O., Rocha de Quadros, M., Rodeck, U. (2003) Targeting the epidermal growth factor receptor in cancer: apoptosis takes center stage. *Cancer Res.* 63:1-5.
168. Kelly, J.M., Hynes, M.J. (1985) Transformation of *Aspergillus niger* by the mdS gene of *Aspergillus nidulans*. *EMBO J.* 4:475-479.
169. Kenmochi, N., Kawaguchi, T., Rozen, S., Davis, E., Goodman, N., Hudson, T.J., Tanaka, T., Page, D.C. (1998) A map of 75 human ribosomal protein genes. *Genome Res.* 8:509-523.
170. Keown, W.A., Campbell, C.R., Kucherlapati, R.S. (1990) Methods for introducing DNA into mammalian cells. *Methods Enzymol.* 185:527-537.
171. Kibbe, A.H., ed. (2000) Handbook of Pharmaceutical Excipients. 3rd ed. Pharmaceutical Press.
172. Kirkpatrick, K.L., Mokbel, K. (2001) The significance of human telomerase reverse transcriptase (hTERT) in cancer. *Eur. J. Surg. Oncol.* 27:754-760.
173. Kirsch, K.H., Georgescu, M.M., Ishimaru, S., Hanafusa, H. (1999) CMS: an adapter molecule involved in cytoskeletal rearrangements. *Proc. Natl. Acad. Sci.* 96:6211-6216.
174. Kiryu-Seo, S., Sasaki, M., Yokohama, H., Nakagomi, S., Hirayama, T., Aoki, S., Wada, K., Kiyama, H. (2000) Damage-induced neuronal endopeptidase (DINE) is a unique metallopeptidase expressed in response to neuronal damage and activates superoxide scavengers. *Proc. Natl. Acad. Sci.* 97:4345-4350.
175. Klarman, G.J., Hawkins, M.E., Le Grice, S.F. (2002) Uncovering the complexities of retroviral ribonuclease H reveals its potential as a therapeutic target. *AIDS Rev.* 4:183-194.
176. Knutson, K.L., Schiffman, K., Disis, M.L. (2001) Immunization with a HER-2/neu helper peptide vaccine generates HER-2/neu CD8 T-cell immunity in cancer patients. *J. Clin. Invest.* 107:477-484.
177. Kobayashi, M., Takezawa, S., Hara, K., Yu, R.T., Umesono, Y., Agata, K., Taniwaki, M., Yasuda, K., Umesono, K. (1999) Identification of a photoreceptor cell-specific nuclear receptor. *Proc. Natl. Acad. Sci.* 96:4814-4819.
178. Kolonin, M.G., Finley, R.L. Jr. (1998) Targeting cyclin-dependent kinases in *Drosophila* with peptide aptamers. *Proc. Natl. Acad. Sci.* 95:14,266-14,271.

179. Korner, C., Knauer, R., Stephani, U., Marquardt, T., Lehle, L., von Figura, K. (1999) Carbohydrate deficient glycoprotein syndrome type IV: deficiency of dolichyl-P-Man:Man(5)GlcNAc(2)-PP-dolichyl mannosyltransferase. *EMBO J.* 18:6816-6822.
180. Kothapalli, R., Buyuksal, I., Wu, S.Q., Chegini, N., Tabibzadeh, S. (1997) Detection of ebaf, a novel human gene of the transforming growth factor beta superfamily association of gene expression with endometrial bleeding. *J. Clin. Invest.* 99:2342-2350.
181. Kovalenko, O.V., Golub, E.I., Bray-Ward, P., Ward, D.C., Radding, C.M. (1997) A novel nucleic acid-binding protein that interacts with human rad51 recombinase. *Nucleic Acids Res.* 25:4946-4953.
182. Kratzschmar, J., Lum, L., Blobel, C.P. (1996) Metargidin, a membrane-anchored metalloprotease-disintegrin protein with an RGD integrin binding sequence. *J. Biol. Chem.* 271:4593-4596.
183. Ku, D.H., Kagan, J., Chen, S.T., Chang, C.D., Baserga, R., Wurzel, J. (1990) The human fibroblast adenine nucleotide translocator gene. Molecular cloning and sequence. *J. Biol. Chem.* 265:16,060-16,063.
184. Kuisle, O., Quiñoá, E., Rigura, R. (1999) Solid phase synthesis of depsides and depsipeptides. *Tetrahedron Lett.* 40:1203-1206.
185. Kunze, G. et al., (1985) Transformation of the industrially important yeasts *Candida maltosa* and *Pichia guilliermondii*. *J. Basic Microbiol.* 25:141-144.
186. Kurtz, M.B., Cortelyou, M.W., Kirsch, D.R. (1986) Integrative transformation of *Candida albicans*, using a cloned *Candida* ADE2 gene. *Mol. Cell. Biol.* 6:142-149.
187. Kyo, S., Takakura, M., Inoue, M. (2000) Telomerase activity in cancer as a diagnostic and therapeutic target. *Histol. Histopathol.* 15:813-824.
188. Lander, E.S. (1999) Array of hope. *Nature Genetics* 21:3-4.
189. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczy, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough,

R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissole, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kasprzyk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer,

- M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., Szustakowki, J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y.J.; International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome *Nature* 409:860-921.
190. Lasham, A., Moloney, S., Hale, T., Homer, C., Zhang, Y.F., Murison, J.G., Braithwaite, A.W., Watson, J. (2003) The Y-box binding protein YB1: A potential negative regulator of the p53 tumor suppressor. *J. Biol. Chem.* Epub ahead of print, June 30, 2003.
191. Lashkari, A., Smith, A.K., Graham, J.M. Jr. (1999) Williams-Beuren syndrome: an update and review for the primary physician. *Clin. Pediatr.* 38:189-208.
192. Lavedan, C. (1998) The synuclein family. *Genome Res.* 8:871-880.
193. Lebacqz-Verheyden, A.M., Kasprzyk, P.G., Raum, M.G., Van Wyke Coelingh, K., Lebacqz, J.A., Battey, J.F. (1988) Posttranslational processing of endogenous and of baculovirus-expressed human gastrin-releasing peptide precursor. *Mol. Cell. Biol.* 8:3129-3135.
194. Lees-Miller, S.P., Anderson, C.W. (1989) Two human 90-kDa heat-shock proteins are phosphorylated in vivo at conserved serines that are phosphorylated *in vitro* by casein kinase II. *J. Biol. Chem.* 264:2431-2437.
195. Lerch, M.M., Gorelick, F.S. (2000) Early trypsinogen activation in acute pancreatitis. *Med. Clin. North Amer.* 84:549-563.
196. Lerner, R.A. (1982) Tapping the immunological repertoire to produce antibodies of predetermined specificity. *Nature* 299:592-596.
197. Li, E., Bestagno, M., Burrone, O. (1996) Molecular cloning and characterization of a transmembrane surface antigen in human cells. *Eur. J. Biochem.* 238:631-638.
198. Lim, D., Orlova, M., Goff, S.P. (Aug. 2002) Mutations of the RNase H C helix of the Moloney murine leukemia virus reverse transcriptase reveal defects in polypurine tract recognition. *J. Virol.* 76:8360-8373.
199. Lin, B., Rommens, J.M., Graham, R.K., Kalchman, M., MacDonald, H., Nasir, J., Delaney, A., Goldberg, Y.P., Hayden, M.R. (1993) Differential 3' polyadenylation of the Huntington disease gene results in two mRNA species with variable tissue expression. *Hum. Mol. Genet.* 2:1541-1545.

200. Lin, W.J., Gary, J.D., Yang, M.C., Clarke, S., Herschman, H.R. (1996) The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J. Biol. Chem.* 271:15,034-15,044.
201. Lin, X., Sikkink, R.A., Rusnak, F., Barber, D.L. (1999) Inhibition of calcineurin phosphatase activity by a calcineurin B homologous protein. *J. Biol. Chem.* 274:36,125-36,131.
202. Linnenbach, A.J., Seng, B.A., Wu, S., Robbins, S., Scollon, M., Pyrc, J.J., Druck, T., Huebner, K. (1993) Retroposition in a family of carcinoma-associated antigen genes. *Mol. Cell Biol.* 13:1507-1515.
203. Linstedt, A.D., Hauri, H.P. (1993) Giantin, a novel conserved Golgi membrane protein containing a cytoplasmic domain of at least 350 kDa. *Mol. Biol. Cell* 4:679-693.
204. Lipshutz, R.J., Fodor, S.P.A., Gingeras, T.R., Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nature Genetics* 21:20-24.
205. Liu A.Y., Robinson R.R., Hellstrom K.E., Murray E.D. Jr., Chang C.P., Hellstrom I. (1987a) Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells. *Proc. Natl. Acad. Sci.* 84:3439-3443.
206. Liu, A.Y., Robinson, R.R., Murray, E.D. Jr., Ledbetter, J.A., Hellstrom, I., Hellstrom, K.E. (1987b) Production of a mouse-human chimeric monoclonal antibody to CD20 with potent Fc-dependent biologic activity. *J. Immunol.* 139:3521-3526.
207. Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darness, J. (1999) Molecular Cell Biology. 4th ed. W H Freeman & Co.
208. Loeffen, J.L., Triepels, R.H., van den Heuvel, L.P., Schuelke, M., Buskens, C.A., Smeets, R.J., Trijbels, J.M., Smeitink, J.A. (1998) cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: human complex I cDNA characterization completed. *Biochem. Biophys. Res. Commun.* 253:415-422.
209. Los, M., Burek, C.J., Stroh, C., Benedyk, K., Hug, H., Mackiewicz. (2003) Anticancer drugs of tomorrow: apoptotic pathways as targets for drug design. *Drug Discov. Today* 15:67-77.
210. Lovering R, Trowsdale J. (1991) A gene encoding 22 highly related zinc fingers is expressed in lymphoid cell lines. *Nucleic Acids Res.* 19:2921-2928.

211. Luckow, V., Summers, M. (1988) Trends in the development of baculovirus expression vectors. *Bio/Technology* 6:47-55.
212. MacBeath, G., Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* 289:1760-1763.
213. Machesky, L.M., Reeves, E., Wientjes, F., Mattheyse, F.J., Grogan, A., Totty, N.F., Burlingame, A.L., Hsuan, J.J., Segal, A.W. (1999) Mammalian actin-related protein 2/3 complex localizes to regions of lamellipodial protrusion and is composed of evolutionarily conserved proteins. *Biochem. J.* 328:105-112.
214. Machiels, J.P., van Baren, N., Marchand, M. (2002) Peptide-based cancer vaccines. *Semin. Oncol.* 29:494-502.
215. Mackay, A., Jones, C., Dexter, T., Silva, R.L., Bulmer, K., Jones, A., Simpson, P., Harris, R.A., Jat, P.S., Neville, A.M., Reis, L.F., Lakhani, S.R., O'Hare, M.J. (2003) cDNA microarray analysis of genes associated with ERBB2 (HER2/neu) overexpression in human mammary luminal epithelial cells. *Oncogene* 22:2680-2688.
216. Maeda, S., Kawai, T., Obinata, M., Fujiwara, H., Horiuchi, T., Saeki, Y., Sato, Y., Furusawa, M. (1985) Production of human alpha-interferon in silkworm using a baculovirus vector. *Nature* 315:592-594.
217. Mahajan, M.A., Murray, A., Samuels, H.H. (2002) NRC-interacting factor 1 is a novel cotransducer that interacts with and regulates the activity of the nuclear hormone receptor coactivator NRC. *Mol. Cell Biol.* 22:6883-6894.
218. Mahimkar, R.M., Baricos, W.H., Visaya, O., Pollock, A.S., Lovett, D.H. (2000) Identification, cellular distribution and potential function of the metalloprotease-disintegrin MDC9 in the kidney. *J. Am. Soc. Nephrol.*, 11:595-603.
219. Mahnensmith, R.L., Aronson, P.S. (1985) Interrelationships among quinidine, amiloride, and lithium as inhibitors of the renal Na⁺-H⁺ exchanger. *J. Biol. Chem.* 260:12,586-12,592.
220. Mangi, A.A., Noiseux, N., Kong, D., He, H., Rezvani, M., Ingwall, J.S., Dzau, V.J. (2003) Mesenchymal stem cells modified with Akt prevent remodeling and restore performance of infarcted hearts. *Nat. Med.* 9:1195-1201. Epublished August 10, 2003.
221. Manning, G., Whyte, D.B., Martinez, R., Hunter, T., Sudarsanam, S. (2002) The protein kinase complement of the human genome. *Science* 298:1912-1934.

222. Marotti, K.R., Tomich, C.S. (1989) Simple and efficient oligonucleotide-directed mutagenesis using one primer and circular plasmid DNA template. (1989) *Gene Anal. Tech.* 6:67-70.
223. Martel-Pelletier, J., Welsch, D.J., and Pelletier, J.P. (2001) Metalloproteases and inhibitors in arthritic diseases. *Best Pract. Res. Clin. Rheumatol.* 15:805-829.
224. Martin, B.M., Tsuji, S., LaMarca, M.E., Maysak, K., Eliason, W., Ginns, E.I. (1988) Glycosylation and processing of high levels of active human glucocerebrosidase in invertebrate cells using a baculovirus expression vector. *DNA* 7:99-106.
225. Massari, M.E., Rivera, R.R., Volland, J.R., Quong, M.W., Breit, T.M., van Dongen, J.J., de Smit, O., Murre, C. (1998) Characterization of ABF-1, a novel basic helix-loop-helix transcription factor expressed in activated B lymphocytes. *Mol. Cell Biol.* 18:3130-3139.
226. Matz, M.V., Fradkov, A.F., Labas, Y.A., Savitsky, A.P., Zaraisky, A.G., Markelov, M.L., Lukyanov, S.A. (1999) Fluorescent proteins from nonbioluminescent *Anthozoa* species. *Nat. Biotechnol.* 17:969-973.
227. Mayer, B.J. (2001) SH3 domains: complexity in moderation. *J. Cell Sci.* 114:1253-1263.
228. Mayer, T.U., Kapoor, T.M., Haggarty, S.J., King, R.W., Schreiber, S.L., Mitchison, T.J. (1999) Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286:971-974.
229. McGraw, R.A. III (1984) Dideoxy DNA sequencing with end-labeled oligonucleotide primers. *Anal. Biochem.* 143:298-303.
230. McKusick, V.A.. (2003) OMIM: Online Mendelian Inheritance in Man <http://www.ncbi.nlm.nih.gov>, #104300.
231. McPherson, M.J., Møller, S.G., Benyon, R., Howe, C. (2000) PCR Basics: From Background to Bench. Springer Verlag.
232. Merla, G., Ucla, C., Guipponi, M., Reymond, A. (2002) Identification of additional transcripts in the Williams-Beuren syndrome critical region. *Hum. Genet.* 110:429-438.
233. Miki, H., Setou, M., Kaneshiro, K., Hirokawa, N. (2001) All kinesin superfamily protein, KIF, genes in mouse and human. *Proc. Natl. Acad. Sci.* 98:7004-7011.

234. Milam, A.H., Rose, L., Cideciyan, A.V., Barakat, M.R., Tang, W.X., Gupta, N., Aleman, T.S., Wright, A.F., Stone, E.M., Sheffield, V.C., Jacobson, S.G. (2002) The nuclear receptor NR2E3 plays a role in human retinal photoreceptor differentiation and degeneration. *Proc. Natl. Acad. Sci.* 99:473-478.
235. Milligan, J.F., Matteucci, M.D., Martin, J.C. (1993) Current concepts in antisense drug design. *J. Med. Chem.* 36:1923-1937.
236. Mitch, W.E., Goldberg, A.L. (1996) Mechanisms of muscle wasting. The role of the ubiquitin-proteasome pathway. *N. Engl. J. Med.* 335:1897-1905.
237. Mitchell, D.A., Nair, S.K. (2000) RNA-transfected dendritic cells in cancer immunotherapy. *J. Clin. Invest.* 106:1065-1069.
238. Miyajima A. (2002) Functional analysis of yeast homologue gene associated with human DNA helicase causative syndromes. *Kokuritsu Iyakuhin Shokuhin Eisei Kenkyusho Hokoku* 120:53-74.
239. Miyajima, A., Schreurs, J., Otsu, K., Kondo, A., Arai, K., Maeda, S. (1987) Use of the silkworm, *Bombyx mori*, and an insect baculovirus vector for high-level expression and secretion of biologically active mouse interleukin-3. *Gene* 58:273-281.
240. Monfardini, C., Schiavon, O., Caliceti, P., Morpurgo, M., Harris, J.M., Veronese, F.M. (1995) A branched monomethoxypoly(ethylene glycol) for protein modification. *Bioconjugate Chem.* 6:62-69.
241. Mori, N. (1997) Neuronal growth-associated proteins in neural plasticity and brain aging. *Nihon Shinkei Seishin Yakurigaku Zasshi* 17:159-167.
242. Mortlock, D.P., Nelson, M.R., Innis, J.W. (1996) An efficient method for isolating putative promoters and 5' transcribed sequences from large genomic clones. *Genome Res.* 6:327-335.
243. Murphy, D., Carter, D.A., eds. (1993) Transgenesis Techniques: Principles and Protocols. Humana Press.
244. Myers, E.W., Miller, W. (1988) Optimal alignments in linear space. *Comput. Appl. Biosci.* 4:11-7.
245. Nagata, K., Kawase, H., Handa, H., Yano, K., Yamasaki, M., Ishimi, Y., Okuda, A., Kikuchi, A., Matsumoto, K. (1995) Replication factor encoded by a putative oncogene, set, associated with myeloid leukemogenesis. *Proc. Natl. Acad. Sci.* 92:4279-4283.

246. Naora, H. (1999) Involvement of ribosomal proteins in regulating cell growth and apoptosis: translational modulation or recruitment for extraribosomal activity? *Immunol. Cell Biol.* 77:197-205.
247. Needleman, S.B., Wunch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48:443-453.
248. Nelson, N., Harvey, W.R. (1999) Vacuolar and plasma membrane proton-adenosine triphosphatases. *Physiol. Rev.* 79:361-385.
249. Nishiyama, H., Higashitsuji, H., Yokoi, H., Itoh, K., Danno, S., Matsuda, T., Fujita, J. (1997) Cloning and characterization of human CIRP (cold-inducible RNA-binding protein) cDNA and chromosomal assignment of the gene. *Gene* 204:115-120.
250. Noma, T., Fujisawa, K., Yamashiro, Y., Shinohara, M., Nakazawa, A., Gondo, T., Ishihara, T., Yoshinobu, K. (2001) Structure and expression of human mitochondrial adenylate kinase targeted to the mitochondrial matrix. *Biochem. J.* 358:225-232.
251. Notredame, C., Higgins, D., Heringa, J. (2000) T-Coffee: A novel method for multiple sequence alignments. *J. Molec. Biol.* 302:205-217.
252. Okayama, H., Berg, P. (1983) A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. *Mol. Cell. Biol.* 3:280-289.
253. Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasaki, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, IV, Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A.,

- Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E., Hayashizaki, Y.; FANTOM Consortium; RIKEN Genome Exploration Research Group Phase I & II Team. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* 420:563-573.
254. Oksenberg, J.R., Barcellos, L.F., Hauser, S.L. (1999) Genetic aspects of multiple sclerosis. *Semin. Neurol.* 19:281-288.
 255. Oliver, C.J., Shenolikar, S. (1998) Physiologic importance of protein phosphatase inhibitors. *Frontiers in Bioscience* 3:961-972.
 256. O'Neil, N.J., Martin, R.L., Tomlinson, M.L., Jones, M.R., Coulson, A., Kuwabara, P.E. (2001) RNA-mediated interference as a tool for identifying drug targets. *Am. J. Pharmacogenomics* 1:45-53.
 257. O'Neill, L.A. (2002) Signal transduction pathways activated by the IL-1 receptor/toll-like receptor superfamily. *Curr. Top. Microbiol. Immunol.* 270:47-61.
 258. Page, D.C., Silber, S., Brown, L.G. (1999) Men with infertility caused by AZFc deletion can produce sons by intracytoplasmic sperm injection, but are likely to transmit the deletion and infertility. *Hum. Reprod.* 14:1722-1726.
 259. Pan, C.X., Koeneman, K.S. (1999) A novel tumor-specific gene therapy for bladder cancer. *Med. Hypothesis* 53:130-135.
 260. Pang, T., Wakabayashi, S., Shigekawa, M. (2001) Calcineurin homologous protein as an essential cofactor for Na⁺/H⁺ exchangers. *J. Biol. Chem* 276:17,367-17,372.
 261. Pang, T., Wakabayashi, S., Shigekawa, M. (2002) Expression of calcineurin B homologous protein 2 protects serum deprivation-induced cell death by serum-independent activation of Na⁺/H⁺ exchanger. *J. Biol. Chem.* 277:43,771-43,777.

262. Papagerakis, S., Shabana, A.H., Depondt, J., Gehanno, P., Forest, N. (2003) Immunohistochemical localization of plakophilins (PKP1, PKP2, PKP3, and p0071) in primary oropharyngeal tumors: correlation with clinical parameters. *Hum. Pathol.* 34:565-572.
263. Pearson, W.R. (2000) Flexible sequence similarity searching with the FASTA3 program package. *Methods Mol. Biol.* 132:185-219.
264. Peattie, D.A., Harding, M.W., Fleming, M.A., DeCenzo, M.T., Lippke, J.A., Livingston, D.J., Benasutti, M. (1992) Expression and characterization of human FKBP52, an immunophilin that associates with the 90-kDa heat-shock protein and is a component of steroid receptor complexes. *Proc. Natl. Acad. Sci.* 89:10,974-10,978.
265. Peelle, B., Gururaja, T.L., Payan, D.G., Anderson, D.C. (2001) Characterization and use of green fluorescent proteins from *Renilla mulleri* and *Ptilosarcus guernyi* for the human cell display of functional peptides. *J. Protein Chem.* 20:507-519.
266. Peng, H., Huard, J. (2003) Stem cells in the treatment of muscle and connective tissue diseases. *Curr. Opin. Pharmacol.* 3:329-333.
267. Pepin, K., Momose, F., Ishida, N., Nagata, K. (2001) Molecular cloning of horse Hsp90 cDNA and its comparative analysis with other vertebrate Hsp90 sequences. *J. Vet. Med. Sci.* 63:115-124.
268. Perez Calvo, J.I., Inigo Gil, P., Giraldo Castellano, P., Torralba Cabeza, M.A., Civeira, F., Lario Garcia, S., Pocovi, M., Lara Garcia, S. (2000) Transforming growth factor beta (TGF-beta) in Gaucher's disease. Preliminary results in a group of patients and their carrier and non-carrier relatives *Med. Clin. (Barc)* 115:601-604.
269. Perron, H., Garson, J.A., Bedin, F., Beseme, F., Paranhos-Baccala, G., Komurian-Pradel, F., Mallet, F., Tuke, P.W., Voisset, C., Blond, J.L., Lalande, B., Seigneurin, J.M., Mandrand, B., The Collaborative Research Group on Multiple Sclerosis (1997) Molecular identification of a novel retrovirus repeatedly isolated from patients with multiple sclerosis. *Proc. Natl. Acad. Sci.* 94:7583-7588.
270. Perry, A.C., Jones, R., Hall, L. (1995) Analysis of transcripts encoding novel members of the mammalian metalloprotease-like, disintegrin-like, cysteine-rich (MDC) protein family and their expression in reproductive and non-reproductive monkey tissues. *Biochem. J.* 312(Pt 1):239-244.

271. Pertl, U., Wodrich, H., Ruelmann, J.M., Gillies, S.D., Lode, H.N., Reisfeld, R.A. (2003) Immunotherapy with a posttranscriptionally modified DNA vaccine induces complete protection against metastatic neuroblastoma. *Blood* 101:649-654.
272. Pfutzer, R.H., Whitcomb, D.C. (2001) SPINK1 mutations are associated with multiple phenotypes. *Pancreatology* 1:457-460.
273. Phillips, M.I., ed. (1999a) Antisense Technology, Part A. Methods in Enzymology Vol. 313. Academic Press, Inc.
274. Phillips, M.I., ed. (1999b) Antisense Technology, Part B. Methods in Enzymology Vol. 314. Academic Press, Inc.
275. Pietu, G., Alibert, O., Guichard, V., Lamy, B., Bois, F., Mariage-Sampson, R., Hougatte, R., Soularue, P., Auffray, C. (1996) Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high density cDNA array. *Genome Res.* 6:492-503.
276. Pinkert, C.A., ed. (1994) Transgenic Animal Technology: A Laboratory Handbook. Academic Press.
277. Pisegna, J.R., Wank, S.A. (1996) Cloning and characterization of the signal transduction of four splice variants of the human pituitary adenylate cyclase activating polypeptide receptor. Evidence for dual coupling to adenylate cyclase and phospholipase C. *J. Biol. Chem.* 271:17,267-17,274.
278. Prentki, P., Krisch, H.M. (1984) *In vitro* insertional mutagenesis with a selectable DNA fragment. *Gene* 29:303-313.
279. Price, N.T., Hall, L., Proud, C.G. (1993) Cloning of cDNA for the beta-subunit of rabbit translation initiation factor-2 using PCR. *Biochim. Biophys. Acta* 1216:170-172.
280. Qin, J., Li, L. (2003) Molecular anatomy of the DNA damage and replication checkpoints. *Radiat. Res.* 159:139-148.
281. Racevskis, J., Dill, A., Stockert, R., Fineberg, S.A. (1996) Cloning of a novel nucleolar guanosine 5'-triphosphate binding protein autoantigen from a breast tumor. *Cell. Growth Differ.* 7:271-280.
282. Ramalho-Santos, M. (2002) "Stemness" *Science* 298:597-600.
283. Raval, P. (1994) Qualitative and quantitative determination of mRNA. *J. Pharmacol. Toxicol. Methods* 32:125-127.

284. Rebbe, N.F., Ware, J., Bertina, R.M., Modrich, P., Stafford, D.W. (1987) Nucleotide sequence of a cDNA for a member of the human 90-kDa heat-shock protein family. *Gene* 53:235-245.
285. Rechid, R., Vingron, M., Argos, P. (1989) A new interactive protein sequence alignment program and comparison of its results with widely used algorithms. *Comput. Appl. Biosci.* 5:107-113.
286. Rehli, M., Krause, S.W., Kreutz, M., Andreessen, R. (1995) Carboxypeptidase M is identical to the MAX.1 antigen and its expression is associated with monocyte to macrophage differentiation. *J. Biol. Chem.* 270:15644-15649.
287. Remington, J.P. (1985) Remington's Pharmaceutical Sciences. 17th ed. Mack Publishing Co.
288. Ribardo, D.A., Peterson, J.W., Chopra, A.K. (2002) Phospholipase A2-activating protein--an important regulatory molecule in modulating cyclooxygenase-2 and tumor necrosis factor production during inflammation. *Indian J. Exp. Biol.* 40:129-138.
289. Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., Markham, A.F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nuc. Acids Res.* 18:2887-2890.
290. Ritter, R.C., Brenner, L.A., Tamura, C.S. (1994) Endogenous CCK and the peripheral neural substrates of intestinal satiety. *Ann. N. Y. Acad. Sci.* 713:255-267.
291. Robertson, H.M. (1996) Members of the pogo superfamily of DNA-mediated transposons in the human genome. *Mol. Gen. Genet.* 252:761-766.
292. Robertson, H.M., Zumpano, K.L. (1997) Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene* 205:203-217.
293. Roepman, R., Bernoud-Hubac, N., Schick, D.E., Maugeri, A., Berger, W., Ropers, H.H., Cremers, F.P., Ferreira, P.A. (2000) The retinitis pigmentosa GTPase regulator (RPGR) interacts with novel transport-like proteins in the outer segments of rod photoreceptors. *Hum. Mol. Genet.* 9:2095-2105.
294. Roessler, B.J., Nosal, J.M., Smith, P.R., Heidler, S.A., Palella, T.D., Switzer, R.L., Becker, M.A. (1993) Human X-linked phosphoribosylpyrophosphate synthetase superactivity is associated with distinct point mutations in the PRPS1 gene. *J. Biol. Chem.* 268:26476-26481.

295. Roggenkamp, R., Janowicz, Z., Stanikowski, B., Hollenberg, C.P. (1984) Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast *Hansenula polymorpha*. *Mol. Gen. Genet.* 194:489-493.
296. Rosato, R.R., Grant, S. (2003) Histone deacetylase inhibitors in cancer therapy. *Cancer Biol. Ther.* 2:30-37.
297. Rosen, R.C., McKenna, K.E. (2002) PDE-5 inhibition and sexual response: pharmacological mechanisms and clinical outcomes. *Ann. Rev. Sex Res.* 13:36-88.
298. Rosenblum, M.G., Donato, N.J. (1989) Tumor necrosis factor alpha: a multifaceted peptide hormone. *Crit. Rev. Immunol.* 9:21-44.
299. Rowland, J.M. (2002) Molecular genetic diagnosis of pediatric cancer: current and emerging methods. *Pediatr. Clin. North Am.* 49:1415-1435.
300. Saha, S., Bardelli, A., Buckhaults, P., Velculescu, V.E., Rago, C., St Croix, B., Romans, K.E., Choti, M.A., Lengauer, C., Kinzler, K.W., Vogelstein, B. (2001) A phosphatase associated with metastasis of colorectal cancer. *Science* 294:1343-1346.
301. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.
302. Sambrook, J., Russell, D.W., Sambrook, J. (1989) Molecular Cloning, A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press.
303. Sanchez, E.R., Faber, L.E., Henzel, W.J., Pratt, W.B. (1990) The 56-59-kilodalton protein identified in untransformed steroid receptor complexes is a unique protein that exists in cytosol in a complex with both the 70- and 90-kilodalton heat-shock proteins. *Biochemistry* 29:5145-5152.
304. Sayers, J.R., Krekel, C., Eckstein, F. (1992) Rapid high-efficiency site-directed mutagenesis by the phosphothioate approach. *Biotechniques* 13:592-596.
305. Schaeferling, M., Schiller, S., Paul, H., Kruschina, M., Pavlickova, M., Meerkamp, M., Giammasi, C., Kambhampati, D. (2002) Application of self-assembly techniques in the design of biocompatible protein microarray surfaces. *Electrophoresis* 23:3097-3105.
306. Schaffer, J.E., Lodish, H.F. (1994) Expression cloning and characterization of a novel adipocyte long chain fatty acid transport protein. *Cell* 79:393-395.

307. Schena, M., ed. (1999) DNA Microarrays: A Practical Approach. Oxford Univ. Press.
308. Schena, M., ed. (2000) Microarray Biochip Technology. 1st ed. Eaton Publishing Co.
309. Schlesinger, D.H. (1988a) MacRomolecular Sequencing and Synthesis: Selected Methods and Applications. Wiley-Liss.
310. Schlesinger, D.H., ed. (1988b) Current Methods in Sequence Comparison and Analysis, Macromolecule Sequencing and Synthesis, Selected Methods and Applications, pp. 127-149, Alan R. Liss, Inc.
311. Schonthal, A.H. (2001) Role of serine/threonine protein phosphatase 2A in cancer. *Cancer Lett.* 170:1-13.
312. Seelig, H.P., Schranz, P., Schroter, H., Wiemann, C., Renz, M. (1994) Macroglolin--a new 376 kD Golgi complex outer membrane protein as target of antibodies in patients with rheumatic diseases and HIV infections. *J. Autoimmun.* 7:67-91.
313. Selkoe, D.J. (2001) Presenilin, Notch, and the genesis and treatment of Alzheimer's disease. *Proc. Natl. Acad. Sci.* 98:11,039-11,041.
314. Setlow, J., Hollaender, A., eds. (1986) Genetic Engineering: Principles and Methods. Plenum Pub. Corp.
315. Shamay, M., Barak, O., Doitsh, G., Ben-Dor, I., Shaul, Y. (2002) Hepatitis B virus pX interacts with HBXAP, a PHD finger protein to coactivate transcription. *J. Biol. Chem.* 277:9982-9988.
316. Shao, H., Andres, D.A. (2000) A novel RalGEF-like protein, RGL3, as a candidate effector for rit and Ras. *J. Biol. Chem.* 275:26,914-26,924.
317. Sheppard, P., Kindsvogel, W., Xu, W., Henderson, K., Schlutsmeier, S., Whitmore, T.E., Kuestner, R., Garrigues, U., Birks, C., Roraback, J., Ostrander, C., Dong, D., Shin, J., Presnell, S., Fox, B., Haldeman, B., Cooper, E., Taft, D., Gilbert, T., Grant, F.J., Tackett, M., Krivan, W., McKnight, G., Clegg, C., Foster, D., Klucher, K.M. (2003) IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat. Immunol.* 4:63-68.
318. Shinnick, T.M., Sutcliffe, J.G., Green, N., Lerner, R.A. (1983) Synthetic peptide immunogens as vaccines. *Ann. Rev. Microbiol.* 37:425-446.

319. Shorter, J., Beard, M.B., Seemann, J., Dirac-Svejstrup, A.B., Warren, G. (2002) Sequential tethering of Golgins and catalysis of SNAREpin assembly by the vesicle-tethering protein p115. *J. Cell Biol.* 157:45-62.
320. Siebenlist, U., Simpson, R.B., Gilbert, W. (1980) *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* 20:269-281.
321. Siegal, G.J., Agranoff, B.W., Albers, R.W., Fisher, S.K., Uhler, M.D., eds. (1999) Basic Neurochemistry, Molecular, Cellular, and Medical Aspects. 6th ed. Lippencott, Williams & Wilkins.
322. Sladek, R., Bader, J.A., Giguere, V. (1997) The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. *Mol. Cell Biol.* 17:5400-5409.
323. Slavin, S., Or, R., Aker, M., Shapira, M.Y., Panigrahi, S., Symeonidis, A., Cividalli, G., Nagler, A. (2001) Nonmyeloablative stem cell transplantation for the treatment of cancer and life-threatening nonmalignant disorders: past accomplishments and future goals. *Cancer Chemother. Pharmacol.* 48:S79-S84.
324. Smit, A.F., Riggs, A.D. (1996) Tiggers and DNA transposon fossils in the human genome. *Proc. Natl. Acad. Sci.* 93:1443-1448.
325. Smith, G.E., Ju, G., Ericson, B.L., Moschera, J., Lahm, H.W., Chizzonite, R., Summers, M.D. (1985) Modification and secretion of human interleukin 2 produced in insect cells by a baculovirus expression vector. *Proc. Natl. Acad. Sci.* 82:8404-8408.
326. Smith, T.F., Waterman, M.S. (1981) Comparison of biosequences. *Adv. Appl. Math.* 2:482-489.
327. Soares, M.B. (1997) Identification and cloning of differentially expressed genes. *Curr. Opin. Biotechnol.* 8:542-546.
328. Soejima, H., Kawamoto, S., Akai, J., Miyoshi, O., Arai, Y., Morohka, T., Matsuo, S., Niikawa, N., Kimura, A., Okubo, K., Mukai, T. (2001) Isolation of novel heart-specific genes using the BodyMap database. *Genomics.* 74:115-120.
329. Soulier, S., Vilotte, J.L., L'Huillier, P.J., Mercier, J.C. (1996) Developmental regulation of murine integrin beta 1 subunit- and Hsc73-encoding genes in mammary gland: sequence of a new mouse Hsc73 cDNA. *Gene* 172:285-289.
330. Southern, E., Mir, K., Shchepinov, M. (1999) Molecular interactions on microarrays. *Nature Genetics* 21:5-9.

331. Stein, C.A., Kreig, A.M., eds. (1998) Applied Antisense Oligonucleotide Technology. Wiley-Liss.
332. Steinhaur, C., Wingren, C., Hager, A.C., Borrebaeck, C.A. (2002) Single framework recombinant antibody fragments designed for protein chip applications. *Biotechniques, Supp.*:38-45.
333. Stetler-Stevenson, W.G., Liotta, L.A., Kleiner, D.E. Jr. (1993) Extracellular matrix 6: role of matrix metalloproteinases in tumor invasion and metastasis. *FASEB J.* 7:1434-1441.
334. Stewart, Z.A., Westfall, M.D., Pietenpol, J.A. (2003) Cell-cycle dysregulation and anticancer therapy. *Trends Pharmacol. Sci.* 24:139-145.
335. Stolz, L.E., Tuan, R.S. (1996) Hybridization of biotinylated oligo(dT) for eukaryotic mRNA quantitation. *Mol. Biotechnol.* 6:225-230.
336. Sturm, A., Dignass, A.U. (2002) Modulation of gastrointestinal wound repair and inflammation by phospholipids. *Biochim. Biophys. Acta* 1582:282-288.
337. Stutz, F., Bachi, A., Doerks, T., Braun, I.C., Seraphin, B., Wilm, M., Bork, P., Izaurralde, E. (2000) REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* 6:638-650.
338. Suh, Y.H., Checler, F. (2002) Amyloid precursor protein, presenilins, and alpha-synuclein: molecular pathogenesis and pharmacological applications in Alzheimer's disease. *Pharmacol. Rev.* 54:469-525.
339. Sutcliffe, J.G., Shinnick, T.M., Green, N., Lerner, R.A. (1983) Antibodies that react with predetermined sites on proteins. *Science* 219:660-666.
340. Tan, J., Town, T., Paris, D., Mori, T., Suo, Z., Crawford, F., Mattson, M.P., Flavell, R.A., Mullan, M. (1999) Microglial activation resulting from CD40-CD40L interaction after beta-amyloid stimulation. *Science* 286:2352-2355.
341. Tang, D.C., DeVit, M., Johnston, S.A. (1992) Genetic immunization is a simple method for eliciting an immune response. *Nature* 356:152-154.
342. Tekur, S., Pawlak, A., Guellaen, G., Hecht, N.B. (1999) Contrin, the human homologue of a germ-cell Y-box-binding protein: cloning, expression, and chromosomal localization. *J. Androl.* 20:135-144.
343. Terada, R., Yamamoto, K., Hakoda, T., Shimada, N., Okano, N., Baba, N., Ninomiya, Y., Gershwin, M.E., Shiratori, Y. (2003) Stromal cell-derived factor-1

- from biliary epithelial cells recruits CXCR4-positive cells: implications for inflammatory liver diseases. *Lab. Invest.* 83:665-672.
344. Thompson, J.D., Higgins, D.G., Gibbon, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-80.
 345. Tilburn, J., Scazzocchio, C., Taylor, G.G., Zabicky-Zissman, J.H., Lockington, R.A., Davies, R.W. (1983) Transformation by integration in *Aspergillus nidulans*. *Gene* 26:205-221.
 346. Trounson, A. (2002) Human embryonic stem cells: mother of all cell and tissue types. *Reprod. Biomed. Online* 4 Suppl. 1:58-63.
 347. Tsuda, T., Gallup, M., Jany, B., Gum, J., Kim, Y., Basbaum, C. (1993) Characterization of a rat airway cDNA encoding a mucin-like protein. *Biochem. Biophys. Res. Commun.* 195:363-373.
 348. Tukey, R.H., Pendurthi, U.R., Nguyen, N.T., Green, M.D., Tephly, T.R. (1993) Cloning and characterization of rabbit liver UDP-glucuronosyltransferase cDNAs. Developmental and inducible expression of 4-hydroxybiphenyl UGT2B13. *J. Biol. Chem.* 268:15,260-15,266.
 349. Turgeman, G., Aslan, H., Gazit, Z., Gazit, D. (2002) Cell-mediated gene therapy for bone formation and regeneration. *Curr. Opin. Mol. Ther.* 4:390-394.
 350. Vainberg, I.E., Lewis, S.A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H.L., Cowan, N.J. (1998) Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell* 93:863-873.
 351. Vale, R.D. (2003) The molecular motor toolbox for intracellular transport. *Cell* 112:467-480.
 352. Vallejo, M., Ron, D., Miller, C.P., Habener, J.F. (1993) C/ATF, a member of the activating transcription factor family of DNA-binding proteins, dimerizes with CAAT/enhancer-binding proteins and directs their binding to cAMP response elements. *Proc. Natl. Acad. Sci.* 90:4679-4683.
 353. Van Damme, A., Vanden Driessche, T., Collen, D., Chuah, M.K. (2002) Bone marrow stromal cells as targets for gene therapy. *Curr. Gene Ther.* 2:195-209.
 354. van den Berg, J.A., van der Laken, K.J., van Ooyen, A.J., Renniers, T.C., Rietveld, K., Schaap, A., Brake, A.J., Bishop, R.J., Schultz, K., Moyer, D. (1990)

- Kluyveromyces* as a host for heterologous gene expression: expression and secretion of prochymosin. *Bio/Technology* 8:135-139.
355. Van den Berghe, L., Laurell, H., Huez, I., Zanibellato, C., Prats, H., Bugler, B. (2000) FIF [fibroblast growth factor-2 (FGF-2)-interacting-factor], a nuclear putatively antiapoptotic factor, interacts specifically with FGF-2. *Mol. Endocrinol.* 14:1709-1724.
 356. Van Den Blink, B., Ten Hove T., Van Den Brink G.R., Peppelenbosch M.P., Van Deventer S.J. (2002) From extracellular to intracellular targets, inhibiting MAP kinases in treatment of Crohn's disease. *Ann. N. Y. Acad. Sci.* 973:349-58.
 357. van der Spoel, A.C., Jeyakumar, M., Butters, T.D., Charlton, H.M., Moore, H.D., Dwek, R.A., Platt, F.M. (2002) Reversible infertility in male mice after oral administration of alkylated imino sugars: a nonhormonal approach to male contraception. *Proc. Natl. Acad. Sci.* 99:17173-17178.
 358. Van Eerdewegh, P., Little, R.D., Dupuis, J., Del Mastro, R.G., Falls, K., Simon, J., Torrey, D., Pandit, S., McKenny, J., Braunschweiger, K., Walsh, A., Liu, Z., Hayward, B., Folz, C., Manning, S.P., Bawa, A., Saracino, L., Thackston, M., Benchekroun, Y., Capparell, N., Wang, M., Adair, R., Feng, Y., Dubois, J., FitzGerald, M.G., Huang, H., Gibson, R., Allen, K.M., Pedan, A., Danzig, M.R., Umland, S.P., Egan, R.W., Cuss, F.M., Rorke, S., Clough, J.B., Holloway, J.W., Holgate, S.T., Keith, T.P. (2002) Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature.* 418:426-430.
 359. Van Laar, J.M., Tyndall, A. (2003) Intense immunosuppression and stem-cell transplantation for patients with severe rheumatic autoimmune disease: a review. *Cancer Control* 10:57-65.
 360. Verhey, K.J., Meyer, D., Deehan, R., Blenis, J., Schnapp, B.J., Rapoport, T.A., Margolis, B. (2001) Cargo of kinesin identified as JIP scaffolding proteins and associated signaling molecules. *J. Cell Biol.* 152:959-970.
 361. Vlak, J.M., Klinkenberg, F.A., Zaal, K.J., Usmany, M., Klinge -Roode, E.C., Geervliet, J.B., Roosien, J., van Lent, J.W. (1988) Functional studies on the p10 gene of Autographa californica nuclear polyhedrosis virus using a recombinant expressing a p10-beta- galactosidase fusion gene. *J. Gen. Virol.* 69:765-776.
 362. Voisset, C., Bouton, O., Bedin, F., Duret, L., Mandrand, B., Mallet, F., Paranhos-Baccala, G. (2000) Chromosomal distribution and coding capacity of the

- human endogenous retrovirus HERV-W family. *AIDS Res. Hum. Retroviruses* 16:731-740.
363. Wagner, R.W., Matteucci, M.D., Grant, D., Huang, T., Froehler, B.C. (1996) Potent and selective inhibition of gene expression by an antisense heptanucleotide. *Nat. Biotechnol.* 14:840-844.
364. Wagner, R.W., Matteucci, M.D., Lewis, J.G., Gutierrez, A.J., Moulds, C., Froehler, B.C. (1993) Antisense gene inhibition by oligonucleotides containing C-5 propyne pyrimidines. *Science* 260:1510-1513.
365. Wakefield, L.M., Yang, Y.A., Dukhanina, O. (2000) Transforming growth factor-beta and breast cancer: Lessons learned from genetically altered mouse models. *Breast Cancer Res.* 2:100-106; Epub February 21, 2000.
366. Walker, J.E., Arizmendi, J.M., Dupuis, A., Fearnley, I.M., Finel, M., Medd, S.M., Pilkington, S.J., Runswick, M.J., Skehel, J.M. (1992) Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from bovine heart mitochondria. Application of a novel strategy for sequencing proteins using the polymerase chain reaction. *J. Mol. Biol.* 226:1051-1072.
367. Walsh, A.C., Feulner, J.A., Reilly, A. (2001) Evidence for functionally significant polymorphism of human glutamate cysteine ligase catalytic subunit: association with glutathione levels and drug resistance in the National Cancer Institute tumor cell line panel. *Toxicol. Sci.* 61:218-223.
368. Wang, J., Kirby, C.E., Herbst, R. (2002) The tyrosine phosphatase PRL-1 localizes to the endoplasmic reticulum and the mitotic spindle and is required for normal mitosis. *J. Biol. Chem.* 277:46659-46668.
369. Wang, M.S., Schinzel, A., Kotzot, D., Balmer, D., Casey, R., Chodirker, B.N., Gyftodimou, J., Petersen, M.B., Lopez-Rangel, E., Robinson, W.P. (1999) Molecular and clinical correlation study of Williams-Beuren syndrome: No evidence of molecular factors in the deletion region or imprinting affecting clinical outcome. *Am. J. Med. Genet.* 86:34-43.
370. Wax, S.D., Rosenfield, C.L., Taubman, M.B. (1994) Identification of a novel growth factor-responsive gene in vascular smooth muscle cells. *J. Biol. Chem.* 269:13,041-13,047.

371. Wei, S., Charmley, P., Concannon, P. (1997) Organization, polymorphism, and expression of the human T-cell receptor AV1 subfamily. *Immunogenetics* 45:405-412.
[0414] Weinberg, J.M., Saini, R. (2003) Biologic therapy for psoriasis: the tumor necrosis factor inhibitors infliximab and etanercept. *Cutis*. 71:25-29.
372. Weiner, H.L., Selkoe, D.J. (2002) Inflammation and therapeutic vaccination in CNS diseases. *Nature* 420:879-884.
373. Weiner, M.P., Felts, K.A., Simcox, T.G., Braman, J.C. (1993) A method for the site-directed mono- and multi- mutagenesis of double-stranded DNA. *Gene* 126:35-41.
374. Weinstein, M.E., Grossman, A., Perle, M.A., Wilmot, P.L., Verma, R.S., Silver, R.T., Arlin, Z., Allen, S.L., Amorosi, E., Waintraub, S.E., et al. (1988) The karyotype of Philadelphia chromosome-negative, bcr rearrangement-positive chronic myeloid leukemia. *Cancer Genet Cytogenet.* 35:223-229.
375. Weishaar, R.E., Cain, M.H., Bristol, J.A. (1985) A new generation of phosphodiesterase inhibitors: multiple molecular forms of phosphodiesterase and the potential for drug selectivity. *J. Med. Chem.* 28:537-545.
376. Weissman, I.L. (2000) Translating stem and progenitor cell biology to the clinic: barriers and opportunities. *Science* 287:1442-1446.
377. Weng, S., Gu, K., Hammond, P.W., Lohse, P., Rise, C., Wagner, R.W., Wright, M.C., Kuimelis, R.G. (2002) Generating addressable protein microarrays with PROfusion covalent mRNA-protein fusion technology. *Proteomics* 2:48-57.
378. Wenger, R.H., Rochelle, J.M., Seldin, M.F., Kohler, G., Nielsen, P.J. (1993) The heat stable antigen (mouse CD24) gene is differentially regulated but has a housekeeping promoter. *J. Biol. Chem.* 268:23,345-23,352.
379. Werner, T., Brack-Werner, R., Leib-Mosch, C., Backhaus, H., Erfle, V., Hehlmann, R. (1990) S71 is a phylogenetically distinct endogenous retroviral element with structural and sequence homology to mimian sarcoma virus (SSV). *Virology* 174:225-238.
380. Wick, G., Kromer, G., Neu, N., Fassler, R., Ziemiecki, A., Muller, R.G., Ginzel, M., Beladi, I., Kuhr, T., Hala, K. (1987) The multi-factorial pathogenesis of autoimmune disease. *Immunol. Lett.* 16:249-257.

381. Wieczorek, H., Brown, D., Grinstein, S., Ehrenfeld, J., Harvey, W.R. (1999) Animal plasma membrane energization by proton-motive V-ATPases. *Bioessays* 21:637-648.
382. Wieser, R. (2002) Rearrangements of chromosomal band 3q21 in myeloid leukemia. *Leuk. Lymphoma* 43:59-65.
383. Winssinger, N., Ficarro, S., Schultz, P.G., and Harris, J.L. (2002) Profiling protein function with small molecule microarrays. *Proc. Natl. Acad. Sci.* 99:11,139-11,144.
384. Wojtowicz-Praga, S. (1999) Clinical potential of matrix metalloprotease inhibitors. *Drugs R. D.* 1:117-129.
385. Wu, A.M., Gallo, R.C. (1975) Reverse Transcriptase. *CRC Crit. Rev. Biochem.* 3:289-347.
386. Xu, C.W., Mendelsohn, A.R., Brent, R. (1997) Cells that register logical relationships among proteins. *Proc. Natl. Acad. Sci. (USA)* 94:12,473-12,478.
387. Xu, Y., Piston, D.W., Johnson, C.H. (1999) A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins. *Proc. Natl. Acad. Sci.* 96:151-156.
388. Yang, N., Shigeta, H., Shi, H., Teng, C.T. (1996) Estrogen-related receptor, hERR1, modulates estrogen receptor-mediated response of human lactoferrin gene promoter. *J. Biol. Chem.* 271:5795-5804.
389. Yelton, M.M., Hamer, J.E., Timberlake, W.E. (1984) Transformation of *Aspergillus nidulans* by using a trpC plasmid. *Proc. Natl. Acad. Sci.* 81:1470-1474.
390. Yoshihama, M., Uechi, T., Asakawa, S., Kawasaki, K., Kato, S., Higa, S., Maeda N., Minoshima, S., Tanaka, T., Shimizu, N., Kenmochi, N. (2002) The human ribosomal protein genes: sequencing and comparative analysis of 73 genes. *Genome Res.* 12:379-390.
391. Yu, L., Zhang, Z., Loewenstein, P.M., Desai, K., Tang, Q., Mao, D., Symington, J.S., Green, M. (1995) Molecular cloning and characterization of a cellular protein that interacts with the human immunodeficiency virus type 1 Tat transactivator and encodes a strong transcriptional activation domain. *J. Virol.* 69:3007-3016.
392. Yu, Z., Restifo, N.P. (2002) Cancer vaccines: progress reveals new complexities. *J. Clin. Invest.* 110:289-294.

393. Zallipsky, S. (1995) Functionalized poly(ethylene glycols) for preparation of biologically relevant conjugates. *Bioconjugate Chem.*, 6:150-165.
394. Zhang, Q., Acland, G.M., Wu, W.X., Johnson, J.L., Pearce-Kelling, S., Tulloch, B., Vervoort, R., Wright, A.F., Aguirre, G.D. (2002) Different RPGR exon ORF15 mutations in Canids provide insights into photoreceptor cell degeneration. *Hum. Mol. Genet.* 11:993-1003.
395. Zhang, W.M., Popova, S.N., Bergman, C., Velling, T., Gullberg, M.K., Gullberg, D. (2002) Analysis of the human integrin alpha11 gene (ITGA11) and its promoter. *Matrix Biol.* 21:513-523.
396. Zhao, H., Grabowski, G.A. (2002) Gaucher disease: Perspectives on a prototype lysosomal disease. *Cell Mol. Life Sci.* 59:694-707.
397. Zhao, N., Hashida, H., Takhshi, N., Misumi, Y., Sakaki, Y. (1995) High-density cDNA filter analysis: a novel approach for large-scale quantitative analysis of gene expression. *Gene* 156:207-215.
398. Zhao, Y., Hong, D.H., Pawlyk, B., Yue, G., Adamian, M., Grynberg, M., Godzik, A., Li, T. (2003) The retinitis pigmentosa GTPase regulator (RPGR)-interacting protein: Subservient RPGR function and participating in disk morphogenesis. *Proc. Natl. Acad. Sci.* 100:3965-3970
399. Zhu, D.L. (1989) Oligonucleotide-directed cleavage and repair of a single stranded vector: a method of site-specific mutagenesis. *Anal. Biochem.* 177:120-124.
400. Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, P., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R.A., Gerstein, M., Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* 293:2101-2105.
401. Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., Snyder, M. (2000) Analysis of yeast protein kinases using protein chips. *Nat. Genetics* 26:283-289.
402. Zhu, H., Snyder, M. (2003) Protein chip technology. *Curr. Opin. Chem. Biol.* 7:55-63.
403. Zhu, J., Kahn, C.R. (1997) Analysis of a peptide hormone-receptor interaction in the yeast two-hybrid system *Proc. Natl. Acad. Sci.* 94:13,063-13,068.

CLAIM

1. A first nucleic acid molecule comprising a polynucleotide sequence chosen from at least one polynucleotide sequence according to SEQ. ID. NOS. 1- 187; SEQ. ID. NOS. 375-484, or a complement thereof, or from at least one polynucleotide sequence that encodes SEQ. ID. NOS. 188-374.

ABSTRACT

The invention provides novel polynucleotides, related polypeptides related nucleic acid and polypeptide compositions corresponding to novel human cDNA clones, and related modulators, such as antibodies and small molecule modulators. The invention also provides methods to make and use these polynucleotides, polypeptides, related compositions, and modulators. These methods include diagnostic, prophylactic and therapeutic applications. The compositions and methods of the invention are useful in treating proliferative disorders, *e.g.*, cancers, and inflammatory, immune, bacterial, and viral disorders.

Table 1. Identification Numbers

FP ID	SEQ.ID.NO. (N1)	SEQ.ID.NO. (P1)	SEQ.ID.NO. (N0)	Clone ID
HG1014903	SEQ.ID.NO. 1	SEQ.ID.NO. 188	SEQ.ID.NO. 375	PLT00014330A02.contig.a
HG1014904	SEQ.ID.NO. 2	SEQ.ID.NO. 189		PLT00014330A02.contig.b
HG1014905	SEQ.ID.NO. 3	SEQ.ID.NO. 190	SEQ.ID.NO. 376	PLT00014330A08.contig.a
HG1014906	SEQ.ID.NO. 4	SEQ.ID.NO. 191		PLT00014330A08.contig.b
HG1014907	SEQ.ID.NO. 5	SEQ.ID.NO. 192	SEQ.ID.NO. 377	PLT00014330A17.contig.a
HG1014908	SEQ.ID.NO. 6	SEQ.ID.NO. 193	SEQ.ID.NO. 378	PLT00014330A20.contig.a
HG1014909	SEQ.ID.NO. 7	SEQ.ID.NO. 194	SEQ.ID.NO. 379	PLT00014330B02.contig.a
HG1014910	SEQ.ID.NO. 8	SEQ.ID.NO. 195		PLT00014330B02.contig.b
HG1014911	SEQ.ID.NO. 9	SEQ.ID.NO. 196	SEQ.ID.NO. 380	PLT00014330B04.contig.a
HG1014912	SEQ.ID.NO. 10	SEQ.ID.NO. 197		PLT00014330B04.contig.b
HG1014913	SEQ.ID.NO. 11	SEQ.ID.NO. 198	SEQ.ID.NO. 381	PLT00014330B05.contig.a
HG1014914	SEQ.ID.NO. 12	SEQ.ID.NO. 199	SEQ.ID.NO. 382	PLT00014330B11.contig.a
HG1014915	SEQ.ID.NO. 13	SEQ.ID.NO. 200	SEQ.ID.NO. 383	PLT00014330B13.contig.a
HG1014916	SEQ.ID.NO. 14	SEQ.ID.NO. 201		PLT00014330B13.contig.b
HG1014917	SEQ.ID.NO. 15	SEQ.ID.NO. 202	SEQ.ID.NO. 384	PLT00014330B18.contig.a
HG1014918	SEQ.ID.NO. 16	SEQ.ID.NO. 203		PLT00014330B18.contig.b
HG1014919	SEQ.ID.NO. 17	SEQ.ID.NO. 204	SEQ.ID.NO. 385	PLT00014330C06.contig.a
HG1014920	SEQ.ID.NO. 18	SEQ.ID.NO. 205		PLT00014330C06.contig.b
HG1014921	SEQ.ID.NO. 19	SEQ.ID.NO. 206	SEQ.ID.NO. 386	PLT00014330C12.contig.a
HG1014922	SEQ.ID.NO. 20	SEQ.ID.NO. 207	SEQ.ID.NO. 387	PLT00014330C14.contig.a
HG1014923	SEQ.ID.NO. 21	SEQ.ID.NO. 208	SEQ.ID.NO. 388	PLT00014330C18.contig.a
HG1014924	SEQ.ID.NO. 22	SEQ.ID.NO. 209		PLT00014330C18.contig.b
HG1014925	SEQ.ID.NO. 23	SEQ.ID.NO. 210	SEQ.ID.NO. 389	PLT00014330D03.contig.a
HG1014926	SEQ.ID.NO. 24	SEQ.ID.NO. 211		PLT00014330D03.contig.b
HG1014927	SEQ.ID.NO. 25	SEQ.ID.NO. 212	SEQ.ID.NO. 390	PLT00014330D05.contig.a
HG1014928	SEQ.ID.NO. 26	SEQ.ID.NO. 213		PLT00014330D05.contig.b
HG1014929	SEQ.ID.NO. 27	SEQ.ID.NO. 214	SEQ.ID.NO. 391	PLT00014330D07.contig.a
HG1014930	SEQ.ID.NO. 28	SEQ.ID.NO. 215	SEQ.ID.NO. 392	PLT00014330D10.contig.a
HG1014931	SEQ.ID.NO. 29	SEQ.ID.NO. 216		PLT00014330D10.contig.b
HG1014932	SEQ.ID.NO. 30	SEQ.ID.NO. 217	SEQ.ID.NO. 393	PLT00014330D12.contig.a
HG1014933	SEQ.ID.NO. 31	SEQ.ID.NO. 218		PLT00014330D12.contig.b
HG1014934	SEQ.ID.NO. 32	SEQ.ID.NO. 219	SEQ.ID.NO. 394	PLT00014330D13.contig.a
HG1014935	SEQ.ID.NO. 33	SEQ.ID.NO. 220	SEQ.ID.NO. 395	PLT00014330D15.contig.a
HG1014936	SEQ.ID.NO. 34	SEQ.ID.NO. 221		PLT00014330D15.contig.b
HG1014937	SEQ.ID.NO. 35	SEQ.ID.NO. 222	SEQ.ID.NO. 396	PLT00014330D17.contig.a
HG1014938	SEQ.ID.NO. 36	SEQ.ID.NO. 223	SEQ.ID.NO. 397	PLT00014330E04.contig.a
HG1014939	SEQ.ID.NO. 37	SEQ.ID.NO. 224	SEQ.ID.NO. 398	PLT00014330E14.contig.a
HG1014940	SEQ.ID.NO. 38	SEQ.ID.NO. 225		PLT00014330E14.contig.b
HG1014941	SEQ.ID.NO. 39	SEQ.ID.NO. 226	SEQ.ID.NO. 399	PLT00014330E24.contig.a
HG1014942	SEQ.ID.NO. 40	SEQ.ID.NO. 227		PLT00014330E24.contig.b
HG1014943	SEQ.ID.NO. 41	SEQ.ID.NO. 228	SEQ.ID.NO. 400	PLT00014330F01.contig.a
HG1014944	SEQ.ID.NO. 42	SEQ.ID.NO. 229	SEQ.ID.NO. 401	PLT00014330F03.contig.a
HG1014945	SEQ.ID.NO. 43	SEQ.ID.NO. 230		PLT00014330F03.contig.b
HG1014946	SEQ.ID.NO. 44	SEQ.ID.NO. 231	SEQ.ID.NO. 402	PLT00014330F04.contig.a
HG1014947	SEQ.ID.NO. 45	SEQ.ID.NO. 232		PLT00014330F04.contig.b
HG1014948	SEQ.ID.NO. 46	SEQ.ID.NO. 233	SEQ.ID.NO. 403	PLT00014330F05.contig.a
HG1014949	SEQ.ID.NO. 47	SEQ.ID.NO. 234	SEQ.ID.NO. 404	PLT00014330F13.contig.a
HG1014950	SEQ.ID.NO. 48	SEQ.ID.NO. 235	SEQ.ID.NO. 405	PLT00014330G21.contig.a
HG1014951	SEQ.ID.NO. 49	SEQ.ID.NO. 236		PLT00014330G21.contig.b

FP ID	SEQ.ID.NO. (N1)	SEQ.ID.NO. (P1)	SEQ.ID.NO. (N0)	Clone ID
HG1014952	SEQ.ID.NO. 50	SEQ.ID.NO. 237		PLT00014330H05.contig.b
HG1014953	SEQ.ID.NO. 51	SEQ.ID.NO. 238	SEQ.ID.NO. 406	PLT00014330H06.contig.a
HG1014954	SEQ.ID.NO. 52	SEQ.ID.NO. 239	SEQ.ID.NO. 407	PLT00014330H12.contig.a
HG1014955	SEQ.ID.NO. 53	SEQ.ID.NO. 240		PLT00014330H12.contig.b
HG1014956	SEQ.ID.NO. 54	SEQ.ID.NO. 241	SEQ.ID.NO. 408	PLT00014330H14.contig.a
HG1014957	SEQ.ID.NO. 55	SEQ.ID.NO. 242		PLT00014330H14.contig.b
HG1014958	SEQ.ID.NO. 56	SEQ.ID.NO. 243	SEQ.ID.NO. 409	PLT00014330H18.contig.a
HG1014959	SEQ.ID.NO. 57	SEQ.ID.NO. 244		PLT00014330H18.contig.b
HG1014960	SEQ.ID.NO. 58	SEQ.ID.NO. 245	SEQ.ID.NO. 410	PLT00014330I11.contig.a
HG1014961	SEQ.ID.NO. 59	SEQ.ID.NO. 246	SEQ.ID.NO. 411	PLT00014330I12.contig.a
HG1014962	SEQ.ID.NO. 60	SEQ.ID.NO. 247		PLT00014330I12.contig.b
HG1014963	SEQ.ID.NO. 61	SEQ.ID.NO. 248	SEQ.ID.NO. 412	PLT00014330I13.contig.a
HG1014964	SEQ.ID.NO. 62	SEQ.ID.NO. 249		PLT00014330I13.contig.b
HG1014965	SEQ.ID.NO. 63	SEQ.ID.NO. 250	SEQ.ID.NO. 413	PLT00014330J10.contig.a
HG1014966	SEQ.ID.NO. 64	SEQ.ID.NO. 251		PLT00014330J10.contig.b
HG1014967	SEQ.ID.NO. 65	SEQ.ID.NO. 252	SEQ.ID.NO. 414	PLT00014330J14.contig.a
HG1014968	SEQ.ID.NO. 66	SEQ.ID.NO. 253		PLT00014330J14.contig.b
HG1014969	SEQ.ID.NO. 67	SEQ.ID.NO. 254	SEQ.ID.NO. 415	PLT00014330J15.contig.a
HG1014970	SEQ.ID.NO. 68	SEQ.ID.NO. 255	SEQ.ID.NO. 416	PLT00014330J21.contig.a
HG1014971	SEQ.ID.NO. 69	SEQ.ID.NO. 256		PLT00014330J21.contig.b
HG1014972	SEQ.ID.NO. 70	SEQ.ID.NO. 257	SEQ.ID.NO. 417	PLT00014330K01.contig.a
HG1014973	SEQ.ID.NO. 71	SEQ.ID.NO. 258	SEQ.ID.NO. 418	PLT00014330K08.contig.a
HG1014974	SEQ.ID.NO. 72	SEQ.ID.NO. 259		PLT00014330K08.contig.b
HG1014975	SEQ.ID.NO. 73	SEQ.ID.NO. 260	SEQ.ID.NO. 419	PLT00014330K09.contig.a
HG1014976	SEQ.ID.NO. 74	SEQ.ID.NO. 261		PLT00014330K09.contig.b
HG1014977	SEQ.ID.NO. 75	SEQ.ID.NO. 262	SEQ.ID.NO. 420	PLT00014330K15.contig.a
HG1014978	SEQ.ID.NO. 76	SEQ.ID.NO. 263		PLT00014330K15.contig.b
HG1014979	SEQ.ID.NO. 77	SEQ.ID.NO. 264	SEQ.ID.NO. 421	PLT00014330K24.contig.a
HG1014980	SEQ.ID.NO. 78	SEQ.ID.NO. 265	SEQ.ID.NO. 422	PLT00014330L01.contig.a
HG1015004	SEQ.ID.NO. 79	SEQ.ID.NO. 266	SEQ.ID.NO. 423	PLT00014330L24.contig.a
HG1014981	SEQ.ID.NO. 80	SEQ.ID.NO. 267	SEQ.ID.NO. 424	PLT00014330M02.contig.a
HG1014982	SEQ.ID.NO. 81	SEQ.ID.NO. 268		PLT00014330M02.contig.b
HG1014983	SEQ.ID.NO. 82	SEQ.ID.NO. 269	SEQ.ID.NO. 425	PLT00014330M08.contig.a
HG1014984	SEQ.ID.NO. 83	SEQ.ID.NO. 270		PLT00014330M08.contig.b
HG1014985	SEQ.ID.NO. 84	SEQ.ID.NO. 271	SEQ.ID.NO. 426	PLT00014330M15.contig.a
HG1014986	SEQ.ID.NO. 85	SEQ.ID.NO. 272	SEQ.ID.NO. 427	PLT00014330M17.contig.a
HG1014987	SEQ.ID.NO. 86	SEQ.ID.NO. 273		PLT00014330M17.contig.b
HG1014988	SEQ.ID.NO. 87	SEQ.ID.NO. 274	SEQ.ID.NO. 428	PLT00014330N10.contig.a
HG1014989	SEQ.ID.NO. 88	SEQ.ID.NO. 275		PLT00014330N10.contig.b
HG1014990	SEQ.ID.NO. 89	SEQ.ID.NO. 276	SEQ.ID.NO. 429	PLT00014330N12.contig.a
HG1014991	SEQ.ID.NO. 90	SEQ.ID.NO. 277		PLT00014330N12.contig.b
HG1014992	SEQ.ID.NO. 91	SEQ.ID.NO. 278	SEQ.ID.NO. 430	PLT00014330N13.contig.a
HG1014993	SEQ.ID.NO. 92	SEQ.ID.NO. 279		PLT00014330N13.contig.b
HG1014994	SEQ.ID.NO. 93	SEQ.ID.NO. 280	SEQ.ID.NO. 431	PLT00014330N22.contig.a
HG1014995	SEQ.ID.NO. 94	SEQ.ID.NO. 281		PLT00014330N22.contig.b
HG1014996	SEQ.ID.NO. 95	SEQ.ID.NO. 282	SEQ.ID.NO. 432	PLT00014330O03.contig.a
HG1014997	SEQ.ID.NO. 96	SEQ.ID.NO. 283	SEQ.ID.NO. 433	PLT00014330O07.contig.a
HG1014998	SEQ.ID.NO. 97	SEQ.ID.NO. 284		PLT00014330O07.contig.b
HG1015005	SEQ.ID.NO. 98	SEQ.ID.NO. 285	SEQ.ID.NO. 434	PLT00014330O18.contig.a
HG1015006	SEQ.ID.NO. 99	SEQ.ID.NO. 286		PLT00014330O18.contig.b
HG1014999	SEQ.ID.NO. 100	SEQ.ID.NO. 287	SEQ.ID.NO. 435	PLT00014330P07.contig.a
HG1015000	SEQ.ID.NO. 101	SEQ.ID.NO. 288		PLT00014330P07.contig.b

FP ID	SEQ.ID.NO. (N1)	SEQ.ID.NO. (P1)	SEQ.ID.NO. (N0)	Clone ID
HG1015001	SEQ.ID.NO. 102	SEQ.ID.NO. 289	SEQ.ID.NO. 436	PLT00014330P09.contig.a
HG1015002	SEQ.ID.NO. 103	SEQ.ID.NO. 290		PLT00014330P09.contig.b
HG1015003	SEQ.ID.NO. 104	SEQ.ID.NO. 291	SEQ.ID.NO. 437	PLT00014330P15.contig.a
HG1015007	SEQ.ID.NO. 105	SEQ.ID.NO. 292	SEQ.ID.NO. 438	PLT00014333A03.contig.a
HG1015008	SEQ.ID.NO. 106	SEQ.ID.NO. 293		PLT00014333A03.contig.b
HG1015009	SEQ.ID.NO. 107	SEQ.ID.NO. 294	SEQ.ID.NO. 439	PLT00014333A06.contig.a
HG1015010	SEQ.ID.NO. 108	SEQ.ID.NO. 295		PLT00014333A06.contig.b
HG1015011	SEQ.ID.NO. 109	SEQ.ID.NO. 296	SEQ.ID.NO. 440	PLT00014333A08.contig.a
HG1015012	SEQ.ID.NO. 110	SEQ.ID.NO. 297	SEQ.ID.NO. 441	PLT00014333A15.contig.a
HG1015013	SEQ.ID.NO. 111	SEQ.ID.NO. 298		PLT00014333A15.contig.b
HG1015014	SEQ.ID.NO. 112	SEQ.ID.NO. 299	SEQ.ID.NO. 442	PLT00014333A16.contig.a
HG1015015	SEQ.ID.NO. 113	SEQ.ID.NO. 300		PLT00014333A16.contig.b
HG1015016	SEQ.ID.NO. 114	SEQ.ID.NO. 301	SEQ.ID.NO. 443	PLT00014333B03.contig.a
HG1015017	SEQ.ID.NO. 115	SEQ.ID.NO. 302		PLT00014333B03.contig.b
HG1015018	SEQ.ID.NO. 116	SEQ.ID.NO. 303	SEQ.ID.NO. 444	PLT00014333B05.contig.a
HG1015019	SEQ.ID.NO. 117	SEQ.ID.NO. 304		PLT00014333B05.contig.b
HG1015020	SEQ.ID.NO. 118	SEQ.ID.NO. 305	SEQ.ID.NO. 445	PLT00014333B15.contig.a
HG1015021	SEQ.ID.NO. 119	SEQ.ID.NO. 306	SEQ.ID.NO. 446	PLT00014333B17.contig.a
HG1015022	SEQ.ID.NO. 120	SEQ.ID.NO. 307		PLT00014333B17.contig.b
HG1015023	SEQ.ID.NO. 121	SEQ.ID.NO. 308	SEQ.ID.NO. 447	PLT00014333C02.contig.a
HG1015024	SEQ.ID.NO. 122	SEQ.ID.NO. 309		PLT00014333C02.contig.b
HG1015025	SEQ.ID.NO. 123	SEQ.ID.NO. 310	SEQ.ID.NO. 448	PLT00014333C10.contig.a
HG1015026	SEQ.ID.NO. 124	SEQ.ID.NO. 311		PLT00014333C10.contig.b
HG1015027	SEQ.ID.NO. 125	SEQ.ID.NO. 312	SEQ.ID.NO. 449	PLT00014333C16.contig.a
HG1015028	SEQ.ID.NO. 126	SEQ.ID.NO. 313		PLT00014333C16.contig.b
HG1015029	SEQ.ID.NO. 127	SEQ.ID.NO. 314	SEQ.ID.NO. 450	PLT00014333C21.contig.a
HG1015030	SEQ.ID.NO. 128	SEQ.ID.NO. 315		PLT00014333C21.contig.b
HG1015031	SEQ.ID.NO. 129	SEQ.ID.NO. 316	SEQ.ID.NO. 451	PLT00014333C24.contig.a
HG1015032	SEQ.ID.NO. 130	SEQ.ID.NO. 317		PLT00014333C24.contig.b
HG1015033	SEQ.ID.NO. 131	SEQ.ID.NO. 318	SEQ.ID.NO. 452	PLT00014333D07.contig.a
HG1015034	SEQ.ID.NO. 132	SEQ.ID.NO. 319		PLT00014333D07.contig.b
HG1015035	SEQ.ID.NO. 133	SEQ.ID.NO. 320	SEQ.ID.NO. 453	PLT00014333D15.contig.a
HG1015036	SEQ.ID.NO. 134	SEQ.ID.NO. 321		PLT00014333D15.contig.b
HG1015037	SEQ.ID.NO. 135	SEQ.ID.NO. 322	SEQ.ID.NO. 454	PLT00014333E01.contig.a
HG1015038	SEQ.ID.NO. 136	SEQ.ID.NO. 323		PLT00014333E01.contig.b
HG1015039	SEQ.ID.NO. 137	SEQ.ID.NO. 324	SEQ.ID.NO. 455	PLT00014333E04.contig.a
HG1015040	SEQ.ID.NO. 138	SEQ.ID.NO. 325	SEQ.ID.NO. 456	PLT00014333E05.contig.a
HG1015041	SEQ.ID.NO. 139	SEQ.ID.NO. 326		PLT00014333E05.contig.b
HG1015042	SEQ.ID.NO. 140	SEQ.ID.NO. 327	SEQ.ID.NO. 457	PLT00014333E14.contig.a
HG1015043	SEQ.ID.NO. 141	SEQ.ID.NO. 328		PLT00014333E14.contig.b
HG1015086	SEQ.ID.NO. 142	SEQ.ID.NO. 329	SEQ.ID.NO. 458	PLT00014333E15.contig.a
HG1015087	SEQ.ID.NO. 143	SEQ.ID.NO. 330		PLT00014333E15.contig.b
HG1015044	SEQ.ID.NO. 144	SEQ.ID.NO. 331	SEQ.ID.NO. 459	PLT00014333E24.contig.b
HG1015045	SEQ.ID.NO. 145	SEQ.ID.NO. 332	SEQ.ID.NO. 460	PLT00014333F07.contig.a
HG1015046	SEQ.ID.NO. 146	SEQ.ID.NO. 333	SEQ.ID.NO. 461	PLT00014333G01.contig.a
HG1015047	SEQ.ID.NO. 147	SEQ.ID.NO. 334	SEQ.ID.NO. 462	PLT00014333G02.contig.a
HG1015048	SEQ.ID.NO. 148	SEQ.ID.NO. 335		PLT00014333G02.contig.b
HG1015088	SEQ.ID.NO. 149	SEQ.ID.NO. 336	SEQ.ID.NO. 463	PLT00014333G09.contig.a
HG1015089	SEQ.ID.NO. 150	SEQ.ID.NO. 337		PLT00014333G09.contig.b
HG1015049	SEQ.ID.NO. 151	SEQ.ID.NO. 338	SEQ.ID.NO. 464	PLT00014333H11.contig.a
HG1015050	SEQ.ID.NO. 152	SEQ.ID.NO. 339	SEQ.ID.NO. 465	PLT00014333H15.contig.a
HG1015051	SEQ.ID.NO. 153	SEQ.ID.NO. 340		PLT00014333H15.contig.b

FP ID	SEQ.ID.NO. (N1)	SEQ.ID.NO. (P1)	SEQ.ID.NO. (N0)	Clone ID
HG1015052	SEQ.ID.NO. 154	SEQ.ID.NO. 341	SEQ.ID.NO. 466	PLT00014333I18.contig.a
HG1015053	SEQ.ID.NO. 155	SEQ.ID.NO. 342		PLT00014333I18.contig.b
HG1015054	SEQ.ID.NO. 156	SEQ.ID.NO. 343	SEQ.ID.NO. 467	PLT00014333I22.contig.a
HG1015055	SEQ.ID.NO. 157	SEQ.ID.NO. 344		PLT00014333I22.contig.b
HG1015056	SEQ.ID.NO. 158	SEQ.ID.NO. 345	SEQ.ID.NO. 468	PLT00014333J01.contig.a
HG1015057	SEQ.ID.NO. 159	SEQ.ID.NO. 346		PLT00014333J01.contig.b
HG1015058	SEQ.ID.NO. 160	SEQ.ID.NO. 347	SEQ.ID.NO. 469	PLT00014333J13.contig.a
HG1015059	SEQ.ID.NO. 161	SEQ.ID.NO. 348		PLT00014333J13.contig.b
HG1015060	SEQ.ID.NO. 162	SEQ.ID.NO. 349	SEQ.ID.NO. 470	PLT00014333J15.contig.a
HG1015061	SEQ.ID.NO. 163	SEQ.ID.NO. 350		PLT00014333J15.contig.b
HG1015062	SEQ.ID.NO. 164	SEQ.ID.NO. 351	SEQ.ID.NO. 471	PLT00014333J17.contig.a
HG1015063	SEQ.ID.NO. 165	SEQ.ID.NO. 352	SEQ.ID.NO. 472	PLT00014333J23.contig.a
HG1015064	SEQ.ID.NO. 166	SEQ.ID.NO. 353		PLT00014333J23.contig.b
HG1015065	SEQ.ID.NO. 167	SEQ.ID.NO. 354	SEQ.ID.NO. 473	PLT00014333K04.contig.a
HG1015066	SEQ.ID.NO. 168	SEQ.ID.NO. 355		PLT00014333K04.contig.b
HG1015067	SEQ.ID.NO. 169	SEQ.ID.NO. 356	SEQ.ID.NO. 474	PLT00014333K08.contig.a
HG1015068	SEQ.ID.NO. 170	SEQ.ID.NO. 357		PLT00014333K08.contig.b
HG1015069	SEQ.ID.NO. 171	SEQ.ID.NO. 358	SEQ.ID.NO. 475	PLT00014333L13.contig.b
HG1015070	SEQ.ID.NO. 172	SEQ.ID.NO. 359	SEQ.ID.NO. 476	PLT00014333M01.contig.a
HG1015071	SEQ.ID.NO. 173	SEQ.ID.NO. 360		PLT00014333M01.contig.b
HG1015072	SEQ.ID.NO. 174	SEQ.ID.NO. 361	SEQ.ID.NO. 477	PLT00014333M02.contig.a
HG1015073	SEQ.ID.NO. 175	SEQ.ID.NO. 362		PLT00014333M02.contig.b
HG1015074	SEQ.ID.NO. 176	SEQ.ID.NO. 363	SEQ.ID.NO. 478	PLT00014333M07.contig.a
HG1015075	SEQ.ID.NO. 177	SEQ.ID.NO. 364		PLT00014333M07.contig.b
HG1015076	SEQ.ID.NO. 178	SEQ.ID.NO. 365	SEQ.ID.NO. 479	PLT00014333M15.contig.a
HG1015077	SEQ.ID.NO. 179	SEQ.ID.NO. 366		PLT00014333M15.contig.b
HG1015078	SEQ.ID.NO. 180	SEQ.ID.NO. 367	SEQ.ID.NO. 480	PLT00014333N05.contig.a
HG1015079	SEQ.ID.NO. 181	SEQ.ID.NO. 368		PLT00014333N05.contig.b
HG1015080	SEQ.ID.NO. 182	SEQ.ID.NO. 369	SEQ.ID.NO. 481	PLT00014333N11.contig.a
HG1015081	SEQ.ID.NO. 183	SEQ.ID.NO. 370		PLT00014333N11.contig.b
HG1015082	SEQ.ID.NO. 184	SEQ.ID.NO. 371	SEQ.ID.NO. 482	PLT00014333O03.contig.a
HG1015083	SEQ.ID.NO. 185	SEQ.ID.NO. 372		PLT00014333O03.contig.b
HG1015084	SEQ.ID.NO. 186	SEQ.ID.NO. 373	SEQ.ID.NO. 483	PLT00014333O10.contig.a
HG1015085	SEQ.ID.NO. 187	SEQ.ID.NO. 374	SEQ.ID.NO. 484	PLT00014333O17.contig.a

Table 2. Structural Characteristics

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1014903	PLT00014330A02.contig.a	89	0	(1-89)			0		(1-89)	no_pfam
HG1014904	PLT00014330A02.contig.b	87	0	(1-87)			0		(1-87)	no_pfam
HG1014905	PLT00014330A08.contig.a	82	0.55	(27-82)		(1-26)	1	(15-37)	(1-14)(38-82)	no_pfam
HG1014906	PLT00014330A08.contig.b	61	0.62	(24-61)		(6-23)	2	(5-27)(31-53)	(1-4)(28-30)(54-61)	no_pfam
HG1014907	PLT00014330A17.contig.a	66	0.11	(1-66)	(39-66)	(11-38)	0		(1-66)	no_pfam
HG1014908	PLT00014330A20.contig.a	54	0.25	(33-54)		(18-32)	0		(1-54)	no_pfam
HG1014909	PLT00014330B02.contig.a	84	0	(1-84)			0		(1-84)	no_pfam
HG1014910	PLT00014330B02.contig.b	73	0.07	(22-73)	(41-73)	(16-40)	0		(1-73)	no_pfam
HG1014911	PLT00014330B04.contig.a	160	0	(1-160)			0		(1-160)	no_pfam
HG1014912	PLT00014330B04.contig.b	108	0.05	(1-108)	(25-108)	(11-24)	0		(1-108)	no_pfam
HG1014913	PLT00014330B05.contig.a	79	0.02	(1-79)			0		(1-79)	no_pfam
HG1014914	PLT00014330B11.contig.a	68	0.23	(15-68)	(26-68)	(1-25)	0		(1-68)	no_pfam
HG1014915	PLT00014330B13.contig.a	55	0.05	(1-55)	(38-55)	(8-37)	0		(1-55)	no_pfam
HG1014916	PLT00014330B13.contig.b	53	0.01	(1-53)	(20-53)	(1-19)	0		(1-53)	no_pfam
HG1014917	PLT00014330B18.contig.a	74	0.7	(22-74)		(2-21)	0		(1-74)	no_pfam
HG1014918	PLT00014330B18.contig.b	53	0.24	(28-53)	(37-53)	(14-36)	0		(1-53)	no_pfam
HG1014919	PLT00014330C06.contig.a	101	0.53	(20-101)	(44-101)	(19-43)	0		(1-101)	no_pfam
HG1014920	PLT00014330C06.contig.b	65	0.01	(1-65)	(18-65)	(1-17)	0		(1-65)	no_pfam
HG1014921	PLT00014330C12.contig.a	68	0.01	(1-68)	(23-68)	(1-22)	0		(1-68)	no_pfam
HG1014922	PLT00014330C14.contig.a	66	0.02	(1-66)			0		(1-66)	no_pfam
HG1014923	PLT00014330C18.contig.a	64	0	(1-64)	(20-64)	(1-19)	0		(1-64)	no_pfam
HG1014924	PLT00014330C18.contig.b	63	0	(1-63)			0		(1-63)	no_pfam
HG1014925	PLT00014330D03.contig.a	132	0.81	(20-132)		(1-19)	0		(1-132)	no_pfam
HG1014926	PLT00014330D03.contig.b	74	0.43	(37-74)		(15-36)	2	(12-31)(46-68)	(1-11)(32-45)(69-74)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1014927	PLT00014330D05.contig.a	60	0.07	(1-60)	(32-60)	(16-31)	0		(1-60)	no_pfam
HG1014928	PLT00014330D05.contig.b	54	0.39	(1-54)	(27-54)	(1-26)	0		(1-54)	no_pfam
HG1014929	PLT00014330D07.contig.a	85	0.03	(4-85)	(1-85)		0		(1-85)	no_pfam
HG1014930	PLT00014330D10.contig.a	79	0.61	(29-79)	(30-79)	(6-29)	0		(1-79)	no_pfam
HG1014931	PLT00014330D10.contig.b	73	0.87	(22-73)	(20-73)	(1-19)	0		(1-73)	no_pfam
HG1014932	PLT00014330D12.contig.a	116	0.01	(1-116)			1	(21-43)	(1-20)(44-116)	no_pfam
HG1014933	PLT00014330D12.contig.b	54	0.24	(24-54)		(1-23)	0		(1-54)	no_pfam
HG1014934	PLT00014330D13.contig.a	60	0	(1-60)			0		(1-60)	no_pfam
HG1014935	PLT00014330D15.contig.a	92	0.01	(1-92)	(21-92)	(6-20)	0		(1-92)	no_pfam
HG1014936	PLT00014330D15.contig.b	89	0.4	(36-89)	(46-89)	(16-45)	1	(12-34)	(1-11)(35-89)	no_pfam
HG1014937	PLT00014330D17.contig.a	96	0.26	(30-96)	(27-96)	(10-26)	0		(1-96)	no_pfam
HG1014938	PLT00014330E04.contig.a	54	0.02	(1-54)			0		(1-54)	no_pfam
HG1014939	PLT00014330E14.contig.a	68	0.02	(1-68)	(19-68)	(1-18)	0		(1-68)	no_pfam
HG1014940	PLT00014330E14.contig.b	61	0	(1-61)	(27-61)	(9-26)	0		(1-61)	no_pfam
HG1014941	PLT00014330E24.contig.a	112	0.01	(1-112)			0		(1-112)	no_pfam
HG1014942	PLT00014330E24.contig.b	62	0.16	(1-62)	(35-62)	(17-34)	1	(15-34)	(1-14)(35-62)	no_pfam
HG1014943	PLT00014330F01.contig.a	77	0	(1-77)			1	(28-45)	(1-27)(46-77)	no_pfam
HG1014944	PLT00014330F03.contig.a	105	0	(1-105)			0		(1-105)	no_pfam
HG1014945	PLT00014330F03.contig.b	71	0.01	(27-71)	(1-71)		0		(1-71)	no_pfam
HG1014946	PLT00014330F04.contig.a	117	0.9	(18-117)	(20-117)	(1-19)	0		(1-117)	no_pfam
HG1014947	PLT00014330F04.contig.b	104	0.09	(25-104)		(1-24)	0		(1-104)	no_pfam
HG1014948	PLT00014330F05.contig.a	50	0.01	(1-50)	(16-50)	(1-15)	0		(1-50)	no_pfam
HG1014949	PLT00014330F13.contig.a	53	0.26	(28-53)		(1-27)	0		(1-53)	no_pfam
HG1014950	PLT00014330G21.contig.a	146	0.16	(28-146)	(29-146)	(6-28)	0		(1-146)	no_pfam
HG1014951	PLT00014330G21.contig.b	53	0.05	(1-53)			1	(20-42)	(1-19)(43-53)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1014952	PLT00014330H05.contig.b	97	0.01	(1-97)	(25-97)	(1-24)	0		(1-97)	rvt
HG1014953	PLT00014330H06.contig.a	50	0.16	(1-50)	(32-50)	(16-31)	0		(1-50)	no_pfam
HG1014954	PLT00014330H12.contig.a	86	0.65	(19-86)		(1-18)	0		(1-86)	no_pfam
HG1014955	PLT00014330H12.contig.b	76	0.03	(1-76)	(19-76)	(1-18)	0		(1-76)	no_pfam
HG1014956	PLT00014330H14.contig.a	68	0.2	(38-68)	(17-68)	(1-16)	0		(1-68)	no_pfam
HG1014957	PLT00014330H14.contig.b	66	0.05	(29-66)	(1-66)		1	(43-62)	(1-42)(63-66)	no_pfam
HG1014958	PLT00014330H18.contig.a	95	0.94	(21-95)	(19-95)	(1-18)	0		(1-95)	no_pfam
HG1014959	PLT00014330H18.contig.b	77	0.01	(38-77)	(1-77)		0		(1-77)	no_pfam
HG1014960	PLT00014330H11.contig.a	62	0.05	(1-62)			1	(31-53)	(1-30)(54-62)	no_pfam
HG1014961	PLT00014330H12.contig.a	88	0.3	(8-88)	(19-88)	(1-18)	0		(1-88)	no_pfam
HG1014962	PLT00014330H12.contig.b	66	0.51	(8-66)	(16-66)	(1-15)	2	(4-26)(43-65)	(1-3)(27-42)(66-66)	no_pfam
HG1014963	PLT00014330H13.contig.a	103	0.04	(1-103)	(41-103)	(17-40)	0		(1-103)	no_pfam
HG1014964	PLT00014330H13.contig.b	84	0.02	(1-84)	(18-84)	(5-17)	0		(1-84)	no_pfam
HG1014965	PLT00014330H10.contig.a	130	0.05	(16-130)	(1-130)		0		(1-130)	no_pfam
HG1014966	PLT00014330H10.contig.b	103	0	(1-103)			0		(1-103)	no_pfam
HG1014967	PLT00014330H14.contig.a	79	0.02	(32-79)	(1-79)		0		(1-79)	no_pfam
HG1014968	PLT00014330H14.contig.b	57	0.03	(1-57)	(23-57)	(1-22)	0		(1-57)	no_pfam
HG1014969	PLT00014330H15.contig.a	68	0.01	(1-68)			0		(1-68)	no_pfam
HG1014970	PLT00014330H21.contig.a	80	0.1	(1-80)	(25-80)	(10-24)	0		(1-80)	no_pfam
HG1014971	PLT00014330H21.contig.b	68	0.08	(1-68)	(22-68)	(1-21)	0		(1-68)	no_pfam
HG1014972	PLT00014330K01.contig.a	73	0	(1-73)			0		(1-73)	no_pfam
HG1014973	PLT00014330K08.contig.a	99	0.16	(1-99)	(26-99)	(1-25)	1	(73-95)	(1-72)(96-99)	no_pfam
HG1014974	PLT00014330K08.contig.b	50	0.26	(1-50)	(18-50)	(1-17)	2	(5-27)(32-49)	(1-4)(28-31)(50-50)	no_pfam
HG1014975	PLT00014330K09.contig.a	100	0.09	(20-100)		(2-19)	0		(1-100)	no_pfam
HG1014976	PLT00014330K09.contig.b	60	0	(1-60)	(23-60)	(11-22)	0		(1-60)	no_pfam
HG1014977	PLT00014330K15.contig.a	72	0.01	(1-72)	(26-72)	(2-25)	0		(1-72)	no_pfam
HG1014978	PLT00014330K15.contig.b	61	0	(1-61)	(33-61)	(9-32)	0		(1-61)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1014979	PLT00014330K24.contig.a	51	0.17	(37-51)	(29-51)	(8-28)	1	(13-35)	(1-12)(36-51)	no_pfam
HG1014980	PLT00014330L01.contig.a	112	0.13	(37-112)	(19-112)	(1-18)	0		(1-112)	no_pfam
HG1014981	PLT00014330M02.contig.a	106	0.01	(1-106)			0		(1-106)	no_pfam
HG1014982	PLT00014330M02.contig.b	88	0.27	(1-88)	(19-88)	(1-18)	0		(1-88)	no_pfam
HG1014983	PLT00014330M08.contig.a	72	0.46	(32-72)		(18-31)	1	(45-67)	(1-44)(68-72)	no_pfam
HG1014984	PLT00014330M08.contig.b	52	0.29	(31-52)		(17-30)	1	(20-42)	(1-19)(43-52)	no_pfam
HG1014985	PLT00014330M15.contig.a	53	0.07	(1-53)	(53-53)	(19-52)	0		(1-53)	no_pfam
HG1014986	PLT00014330M17.contig.a	110	0.13	(1-110)	(21-110)	(1-20)	0		(1-110)	no_pfam
HG1014987	PLT00014330M17.contig.b	82	0.45	(29-82)	(30-82)	(16-29)	0		(1-82)	no_pfam
HG1014988	PLT00014330N10.contig.a	75	0.15	(38-75)		(18-37)	1	(20-42)	(1-19)(43-75)	no_pfam
HG1014989	PLT00014330N10.contig.b	68	0	(1-68)	(22-68)	(1-21)	0		(1-68)	no_pfam
HG1014990	PLT00014330N12.contig.a	56	0	(1-56)	(33-56)	(18-32)	0		(1-56)	no_pfam
HG1014991	PLT00014330N12.contig.b	56	0	(1-56)	(20-56)	(1-19)	0		(1-56)	no_pfam
HG1014992	PLT00014330N13.contig.a	83	0.87	(23-83)	(20-83)	(1-19)	1	(4-26)	(1-3)(27-83)	no_pfam
HG1014993	PLT00014330N13.contig.b	55	0.29	(28-55)	(29-55)	(14-28)	1	(10-32)	(1-9)(33-55)	no_pfam
HG1014994	PLT00014330N22.contig.a	74	0.02	(1-74)	(33-74)	(19-32)	0		(1-74)	no_pfam
HG1014995	PLT00014330N22.contig.b	57	0.12	(1-57)	(20-57)	(1-19)	0		(1-57)	no_pfam
HG1014996	PLT00014330O03.contig.a	70	0.32	(1-70)	(19-70)	(5-18)	1	(7-29)	(1-6)(30-70)	no_pfam
HG1014997	PLT00014330O07.contig.a	78	0	(1-78)			0		(1-78)	no_pfam
HG1014998	PLT00014330O07.contig.b	73	0.06	(1-73)	(33-73)	(19-32)	0		(1-73)	no_pfam
HG1014999	PLT00014330P07.contig.a	85	0.03	(1-85)	(33-85)	(1-32)	0		(1-85)	no_pfam
HG1015000	PLT00014330P07.contig.b	61	0.05	(34-61)	(32-61)	(1-31)	0		(1-61)	no_pfam
HG1015001	PLT00014330P09.contig.a	101	0.17	(1-101)	(33-101)	(13-32)	0		(1-101)	no_pfam
HG1015002	PLT00014330P09.contig.b	98	0.01	(1-98)			0		(1-98)	no_pfam
HG1015003	PLT00014330P15.contig.a	61	0.02	(1-61)			0		(1-61)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1015004	PLT00014330L24.contig.a	50	0.17	(38-50)	(34-50)	(1-33)	0		(1-50)	no_pfam
HG1015005	PLT00014330O18.contig.a	82	0	(1-82)			0		(1-82)	no_pfam
HG1015006	PLT00014330O18.contig.b	66	0	(1-66)			0		(1-66)	no_pfam
HG1015007	PLT00014333A03.contig.a	83	0.08	(1-83)	(39-83)	(19-38)	1	(15-37)	(1-14)(38-83)	no_pfam
HG1015008	PLT00014333A03.contig.b	64	0.1	(30-64)	(29-64)	(11-28)	0		(1-64)	no_pfam
HG1015009	PLT00014333A06.contig.a	153	0.01	(1-153)			0		(1-153)	no_pfam
HG1015010	PLT00014333A06.contig.b	66	0.13	(35-66)	(33-66)	(18-32)	0		(1-66)	no_pfam
HG1015011	PLT00014333A08.contig.a	66	0.26	(1-66)	(22-66)	(1-21)	0		(1-66)	no_pfam
HG1015012	PLT00014333A15.contig.a	136	0.03	(1-136)			0		(1-136)	no_pfam
HG1015013	PLT00014333A15.contig.b	67	0.8	(38-67)	(35-67)	(17-34)	0		(1-67)	no_pfam
HG1015014	PLT00014333A16.contig.a	51	0.02	(1-51)			0		(1-51)	no_pfam
HG1015015	PLT00014333A16.contig.b	50	0.46	(25-50)	(41-50)	(16-40)	0		(1-50)	no_pfam
HG1015016	PLT00014333B03.contig.a	63	0.02	(1-63)			0		(1-63)	no_pfam
HG1015017	PLT00014333B03.contig.b	50	0	(1-50)	(15-50)	(1-14)	0		(1-50)	no_pfam
HG1015018	PLT00014333B05.contig.a	55	0.05	(1-55)			1	(29-51)	(1-28)(52-55)	no_pfam
HG1015019	PLT00014333B05.contig.b	53	0.49	(1-53)	(18-53)	(1-17)	0		(1-53)	no_pfam
HG1015020	PLT00014333B15.contig.a	53	0	(1-53)	(28-53)	(3-27)	0		(1-53)	no_pfam
HG1015021	PLT00014333B17.contig.a	76	0.35	(16-76)		(1-15)	0		(1-76)	no_pfam
HG1015022	PLT00014333B17.contig.b	65	0.01	(1-65)			1	(42-64)	(1-41)(65-65)	no_pfam
HG1015023	PLT00014333C02.contig.a	77	0.03	(1-77)			0		(1-77)	no_pfam
HG1015024	PLT00014333C02.contig.b	51	0.77	(22-51)		(8-21)	1	(12-34)	(1-11)(35-51)	no_pfam
HG1015025	PLT00014333C10.contig.a	99	0.33	(1-99)	(50-99)	(19-49)	0		(1-99)	no_pfam
HG1015026	PLT00014333C10.contig.b	92	0.21	(18-92)	(20-92)	(1-19)	0		(1-92)	no_pfam
HG1015027	PLT00014333C16.contig.a	363	0.04	(1-363)	(15-363)	(1-14)	0		(1-363)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1015028	PLT00014333C16.contig.b	86	0.24	(1-86)	(27-86)	(1-26)	0		(1-86)	no_pfam
HG1015029	PLT00014333C21.contig.a	82	0.49	(1-82)	(49-82)	(19-48)	0		(1-82)	no_pfam
HG1015030	PLT00014333C21.contig.b	77	0.03	(1-77)	(28-77)	(9-27)	0		(1-77)	no_pfam
HG1015031	PLT00014333C24.contig.a	94	0.11	(1-94)	(30-94)	(15-29)	1	(10-32)	(1-9)(33-94)	no_pfam
HG1015032	PLT00014333C24.contig.b	88	0	(1-88)			2	(34-56)(61-78)	(1-33)(57-60)(79-88)	no_pfam
HG1015033	PLT00014333D07.contig.a	73	0.02	(1-73)	(21-73)	(1-20)	0		(1-73)	no_pfam
HG1015034	PLT00014333D07.contig.b	67	0.23	(1-67)	(32-67)	(1-31)	0		(1-67)	no_pfam
HG1015035	PLT00014333D15.contig.a	64	0.11	(32-64)	(31-64)	(16-30)	0		(1-64)	no_pfam
HG1015036	PLT00014333D15.contig.b	62	0.29	(34-62)	(31-62)	(5-30)	2	(13-32)(42-61)	(1-12)(33-41)(62-62)	no_pfam
HG1015037	PLT00014333E01.contig.a	73	0	(36-73)	(1-73)		1	(26-48)	(1-25)(49-73)	no_pfam
HG1015038	PLT00014333E01.contig.b	67	0.51	(35-67)	(26-67)	(8-25)	1	(10-32)	(1-9)(33-67)	no_pfam
HG1015039	PLT00014333E04.contig.a	53	0.01	(1-53)			0		(1-53)	no_pfam
HG1015040	PLT00014333E05.contig.a	66	0.01	(1-66)	(25-66)	(8-24)	0		(1-66)	no_pfam
HG1015041	PLT00014333E05.contig.b	57	0.03	(1-57)	(45-57)	(1-44)	0		(1-57)	no_pfam
HG1015042	PLT00014333E14.contig.a	108	0.01	(1-108)			0		(1-108)	no_pfam
HG1015043	PLT00014333E14.contig.b	61	0.24	(26-61)	(29-61)	(14-28)	0		(1-61)	no_pfam
HG1015044	PLT00014333E24.contig.b	91	0.01	(1-91)	(32-91)	(18-31)	0		(1-91)	Transposase_1
HG1015045	PLT00014333F07.contig.a	52	0	(1-52)	(17-52)	(1-16)	0		(1-52)	no_pfam
HG1015046	PLT00014333G01.contig.a	69	0.24	(1-69)	(33-69)	(14-32)	0		(1-69)	no_pfam
HG1015047	PLT00014333G02.contig.a	77	0.03	(19-77)	(1-77)		0		(1-77)	no_pfam
HG1015048	PLT00014333G02.contig.b	57	0	(1-57)			0		(1-57)	no_pfam
HG1015049	PLT00014333H11.contig.a	95	0.03	(1-95)	(36-95)	(12-35)	0		(1-95)	no_pfam
HG1015050	PLT00014333H15.contig.a	90	0.23	(35-90)		(1-34)	0		(1-90)	no_pfam
HG1015051	PLT00014333H15.contig.b	60	0	(1-60)			0		(1-60)	no_pfam
HG1015052	PLT00014333I18.contig.a	58	0.69	(22-58)	(34-58)	(12-33)	1	(7-29)	(1-6)(30-58)	no_pfam
HG1015053	PLT00014333I18.contig.b	50	0.77	(22-50)		(1-21)	0		(1-50)	no_pfam
HG1015054	PLT00014333I22.contig.a	70	0.08	(1-70)	(19-70)	(1-18)	0		(1-70)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1015055	PLT00014333J22.contig.b	54	0.96	(23-54)	(25-54)	(1-24)	1	(6-28)	(1-5)(29-54)	no_pfam
HG1015056	PLT00014333J01.contig.a	84	0.03	(1-84)	(35-84)	(19-34)	0		(1-84)	no_pfam
HG1015057	PLT00014333J01.contig.b	66	0.08	(32-66)	(33-66)	(1-32)	0		(1-66)	no_pfam
HG1015058	PLT00014333J13.contig.a	106	0.02	(1-106)			1	(46-68)	(1-45)(69-106)	no_pfam
HG1015059	PLT00014333J13.contig.b	93	0.06	(37-93)	(1-93)		0		(1-93)	no_pfam
HG1015060	PLT00014333J15.contig.a	63	0.12	(1-63)	(17-63)	(1-16)	0		(1-63)	no_pfam
HG1015061	PLT00014333J15.contig.b	62	0.18	(1-62)	(22-62)	(7-21)	1	(20-42)	(1-19)(43-62)	no_pfam
HG1015062	PLT00014333J17.contig.a	88	0	(1-88)	(36-88)	(16-35)	0		(1-88)	no_pfam
HG1015063	PLT00014333J23.contig.a	66	0.05	(1-66)	(16-66)	(1-15)	0		(1-66)	no_pfam
HG1015064	PLT00014333J23.contig.b	57	0.33	(1-57)	(31-57)	(14-30)	0		(1-57)	no_pfam
HG1015065	PLT00014333K04.contig.a	131	0.01	(1-131)			0		(1-131)	Gag_p24
HG1015066	PLT00014333K04.contig.b	125	0.14	(1-125)	(19-125)	(1-18)	0		(1-125)	integrase
HG1015067	PLT00014333K08.contig.a	69	0.19	(1-69)	(34-69)	(19-33)	1	(28-50)	(1-27)(51-69)	no_pfam
HG1015068	PLT00014333K08.contig.b	63	0.17	(21-63)		(1-20)	0		(1-63)	no_pfam
HG1015069	PLT00014333L13.contig.b	52	0	(1-52)			0		(1-52)	rnaseH
HG1015070	PLT00014333M01.contig.a	110	0.29	(1-110)	(20-110)	(1-19)	1	(86-108)	(1-85)(109-110)	no_pfam
HG1015071	PLT00014333M01.contig.b	68	0.01	(1-68)	(18-68)	(1-17)	1	(41-63)	(1-40)(64-68)	no_pfam
HG1015072	PLT00014333M02.contig.a	101	0.01	(38-101)	(43-101)	(12-42)	0		(1-101)	no_pfam
HG1015073	PLT00014333M02.contig.b	50	0	(1-50)	(14-50)	(1-13)	0		(1-50)	no_pfam
HG1015074	PLT00014333M07.contig.a	70	0.26	(37-70)	(30-70)	(4-29)	1	(13-35)	(1-12)(36-70)	no_pfam
HG1015075	PLT00014333M07.contig.b	58	0.62	(15-58)	(16-58)	(1-15)	0		(1-58)	no_pfam
HG1015076	PLT00014333M15.contig.a	80	0.04	(1-80)	(42-80)	(18-41)	0		(1-80)	no_pfam
HG1015077	PLT00014333M15.contig.b	54	0.08	(1-54)	(42-54)	(18-41)	0		(1-54)	no_pfam
HG1015078	PLT00014333N05.contig.a	73	0.1	(5-73)	(15-73)	(1-14)	0		(1-73)	no_pfam
HG1015079	PLT00014333N05.contig.b	70	0.45	(35-70)	(39-70)	(5-38)	0		(1-70)	no_pfam

FP ID	Clone ID	Predicted Protein Length	Tree Vote	Mature Protein Coords.	Alternate Mature Protein Coords.	Signal Peptide Coords.	TM	TM Coords.	Non-TM Coords.	Pfam
HG1015080	PLT00014333N11.contig.a	95	0.01	(1-95)	(30-95)	(15-29)	0		(1-95)	no_pfam
HG1015081	PLT00014333N11.contig.b	69	0.03	(9-69)	(22-69)	(5-21)	0		(1-69)	no_pfam
HG1015082	PLT00014333O03.contig.a	72	0.21	(3-72)	(28-72)	(14-27)	0		(1-72)	no_pfam
HG1015083	PLT00014333O03.contig.b	55	0.01	(1-55)	(25-55)	(10-24)	0		(1-55)	no_pfam
HG1015084	PLT00014333O10.contig.a	55	0.06	(4-55)	(15-55)	(1-14)	0		(1-55)	no_pfam
HG1015085	PLT00014333O17.contig.a	71	0.11	(1-71)	(20-71)	(1-19)	0		(1-71)	no_pfam
HG1015086	PLT00014333E15.contig.a	92	0.49	(20-92)		(1-19)	1	(5-27)	(1-4)(28-92)	no_pfam
HG1015087	PLT00014333E15.contig.b	78	0.01	(1-78)			1	(52-71)	(1-51)(72-78)	no_pfam
HG1015088	PLT00014333G09.contig.a	125	0	(1-125)			0		(1-125)	no_pfam
HG1015089	PLT00014333G09.contig.b	63	0.11	(1-63)	(41-63)	(18-40)	0		(1-63)	no_pfam

Table 3. Similarity to Known Sequences

FP ID	Clone ID	Top Hit Accession ID	Top Hit Annotation	Top Hit % ID	Top Human Hit Accession ID	Top Human Hit Annot	Top Human Hit % ID
HG1014903	PLT00014330A02.contig.a	gi 34529187 dbj BAC85656.1	unnamed protein product [Homo sapiens]	59	gi 34529187 dbj BAC85656.1	unnamed protein product [Homo sapiens]	59
HG1014910	PLT00014330B02.contig.b	gi 7770237 gb AAAF69654.1	PRO2822 [Homo sapiens]	76	gi 7770237 gb AAAF69654.1	PRO2822 [Homo sapiens]	76
HG1014914	PLT00014330B11.contig.a	gi 38085361 ref XP_355822.1	similar to RIKEN cDNA 6330419J24 gene [Mus musculus]	80		no_human_hit	
HG1014933	PLT00014330D12.contig.b	gi 8923214 ref NP_060190.1	signal-transducing adaptor protein-2; brk kinase substrate [Homo sapiens] gi 7020193 dbj BAA91028.1 unnamed protein product [Homo sapiens]	57	gi 8923214 ref NP_060190.1	signal-transducing adaptor protein-2; brk kinase substrate [Homo sapiens] gi 7020193 dbj BAA91028.1 unnamed protein product [Homo sapiens]	57
HG1014948	PLT00014330F05.contig.a	gi 34534372 dbj BAC86987.1	unnamed protein product [Homo sapiens]	56	gi 34534372 dbj BAC86987.1	unnamed protein product [Homo sapiens]	56

FP ID	Clone ID	Top Hit Accession ID	Top Hit Annotation	Top Hit % ID	Top Human Hit Accession ID	Top Human Hit Annot	Top Human Hit % ID
HG1014952	PLT00014330H05.contig.b	gi 2981631 dbj BAA25253.1	ORF2 [Canis familiaris]	58	no_human_hit		
HG1014958	PLT00014330H18.contig.a	gi 13310191 gb AAK18189.1	recombinant envelope protein [multiple sclerosis associated retrovirus element]	52	no_human_hit		
HG1014971	PLT00014330J21.contig.b	gi 23503335 ref NP_694983.1	hypothetical protein FLJ25952 [Homo sapiens] gi 21758947 dbj BAC05422.1 unnamed protein product [Homo sapiens]	64	gi 23503335 ref NP_694983.1	hypothetical protein FLJ25952 [Homo sapiens] gi 21758947 dbj BAC05422.1 unnamed protein product [Homo sapiens]	64
HG1014975	PLT00014330K09.contig.a	gi 34528691 dbj BAC85556.1	unnamed protein product [Homo sapiens]	56	gi 34528691 dbj BAC85556.1	unnamed protein product [Homo sapiens]	56
HG1014977	PLT00014330K15.contig.a	gi 34533624 dbj BAC86755.1	unnamed protein product [Homo sapiens]	81	gi 34533624 dbj BAC86755.1	unnamed protein product [Homo sapiens]	81
HG1014983	PLT00014330M08.contig.a	gi 21754422 dbj BAC04501.1	unnamed protein product [Homo sapiens]	55	gi 21754422 dbj BAC04501.1	unnamed protein product [Homo sapiens]	55
HG1014992	PLT00014330N13.contig.a	gi 37182643 gb AAQ89122.1	DRDL5813 [Homo sapiens]	56	gi 37182643 gb AAQ89122.1	DRDL5813 [Homo sapiens]	56

FP ID	Clone ID	Top Hit Accession ID	Top Hit Annotation	Top Hit % ID	Top Human Hit Accession ID	Top Human Hit Annot	Top Human Hit % ID
HG1015030	PLT00014333C21.contig.b	gi 18027736 gb AAL55829 .1	unknown [Homo sapiens]	87	gi 18027736 gb AAL55829.1	unknown [Homo sapiens]	87
HG1015044	PLT00014333E24.contig.b	gi 1698455 g b AAC52011. 1	mariner transposase [Homo sapiens]	79	gi 1698455 gb AAC52011.1	mariner transposase [Homo sapiens]	79
HG1015082	PLT00014333O03.contig.a	gi 21754422 dbj BA C0450 1.1	unnamed protein product [Homo sapiens]	75	gi 21754422 db j BAC04501.1	unnamed protein product [Homo sapiens]	75

Appendix A

SEQ.ID.NO. 1 HG1014903N1 PLT00014330A02.contig.a
 ATGCATGTGCATGTTTCATTGCGAGCACTATTTCACAACAGCAAAGACATAGACTCAACCCAA
 ATGCCCATCAATGATGCATTGGATGAAGAAAATATGGTATATTACACCATGGAATACTA
 TGCAGCCGTAAAGGGAAACAAGATAATGTTCCTTTGCAAGGACGTGGGTGGAGCTGGAAGC
 CATTATCCTCAGCAAACTAACACAGGAACAGAAAACCAACACCATGTTCTCATTAT
 AAGTGGGAGCCCAACAATGAGAACACATAG

SEQ.ID.NO. 2 HG1014904N1 PLT00014330A02.contig.b
 ATGATGCATTGGATGAAGAAAATATGGTATATTTACACCATGGAATACTATGCAGCCGTA
 AAAGGGAACAAGATAATGTTCTTTTGCAAGGACGTGGGTGGAGCTGGAAGCCATTATCCTC
 AGCAAACTAACACAGGAACAGAAAACCAACACCATGTTCTCACTTATAAGTGGGAGC
 CCAACAATGAGAACACATAGACACAGGGAGGGGAAACAACACACACTTGGGTCTGTCAATGG
 GGGTGGAGGGAGGGAGAGCATTAG

SEQ.ID.NO. 3 HG1014905N1 PLT00014330A08.contig.a
 ATGAGGTTTCCAAATTTCTCTGCATCCTCATCAAAACTTGTTATTATCTGTCITTTTTTATT
 TTAGACGTTCTAGAAAGGTGTGGAGTGGTATCTCAATTGTGGTATTGATTGCAATTCCTTA
 AGGACTAATGCTTTTGAACATTTTTTTCATGTGCTTATTGGCCATTGTGATATCTTCTCTG
 GAGAAATGTTTATTCAAATCCTTTGTCCATTTTTTAATTGGATTGTCITTTTTTATTGTTGC
 ATGGTATAA

SEQ.ID.NO. 4HG1014906N1	PLT00014330A08.contig.b
ATGTTTATTCAAATCCTTTGTCCATTTTTAATTGGATTGTCCTTTTTATTGTTGCATGGT	
ATAAGTTCTCTATATATTCTGGACACAAAGTCTCTTATTAGTTATACCATTTGCAAAATATT	
TTCTCCCATTTCTGTGCGTTGTCTTTTAACCTTTTCTAATGTGTCCATTGAAACACAAAAAGT	
TTTTAA	
SEQ.ID.NO. 5HG1014907N1	PLT00014330A17.contig.a
ATGGGAGACAGCCTCTTACTTTGCCAAGAAATGAAAGGATTGGACCGAGCTGGAAAAACCT	
CCTTTACCAGATGCTGACTGGCACTGGTGGTTTTTGTCTCTCGACAGTATCCACAATAGCT	
GACGGCTGGGTGTTTCAGTTTGAAAAATATTTTGTGTGCCCTTCATCTTCACTGCAATTTTGT	
GTAAATTTCTCAAAGATCTGA	
SEQ.ID.NO. 6HG1014908N1	PLT00014330A20.contig.a
ATGACTTACCTTGAAATCTTTATTCAAAGGATATACTATTCAATTTTTCAGTGTGCCACT	
GCCACCACTGCATGTTTTTGGTCAGAGTGCAGTGCAACTAACATCGCCATCCTGTGCTAGGA	
AAAAGATCTGCAGGTTGCTGGATAGTGGAAGTTGCACACATCTAA	
SEQ.ID.NO. 7HG1014909N1	PLT00014330B02.contig.a
ATGCAAGTACATTTTGGAGTGTAACAAAGCCCCACAAATCTTTTCCTGATTTTGTTTACATG	
ACCTTGAATTTATTGAGTGATGAGAGCTCTGTGCAGAAATCAGATTTTCAGAGCTCCTTGGCC	
CAAGCCAGGAAGTGGGCATGCATGACTGACCAAGCCCCAGAGTGGCAGGAGGCCAACCTGC	

TGTGTCCTTATCTTGCTTTCCCATAGCTCAAGAGGACAAGAGCAGTGTTCTAGGAGAAAGCC
AAATTATTTTCCTAA

SEQ.ID.NO. 8 HG1014910N1 PLT00014330B02.contig.b
ATGGAGTCTCACTCTGTTGCCCAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCAGCGCAA
GCTCTGCCCTCCTGGGTTACACACCAATTCTCCACCTCAGCCCTCCTGAGTAGCCGGGACCAC
AGGCGCCCGCCACTACACCCAGCTAATTTTGTATTGTATTTTGTAGTGGAGACGGGTTTT
CACTGTGTAGCCAGGATGGTCTCAATCTCCTGGCCTCGTGA

SEQ.ID.NO. 9 HG1014911N1 PLT00014330B04.contig.a
ATGAGGATCCCTGAGTTGGGGCAGAGAAAGCAGGCAGGGAGAAATGGGGACCTGCGGC
GCTGGGGGCCCAGCTGGTGGAGATCTCCAGCTAGGAGATCCCCAGCACAGCACTTGCCCT
GCCCTCTGCTTGTTCAGGGATGTAACTTGCCCTCCTCCTGCTTCTCCGGGCACTCAGG
GCACCGCCTGGGAGGGTCTTGTAGTGCACACTCAGCATCTCTGTTCCCTGGAGCGCATG
GGGTGTGAAACGCCCTGATGCCAGCCAGCTCCCCAGTCTTGAGAGGCTGGTTGAATGGCAC
ATCTCCCTAGGAGGCTCCCTCCCCAGGGTCCCTCTGCCCCCTGCTGTCCATGCAGTCGGT
CCTCACCCGTCAGGAAACGTAGCCTGCTTGCCCTGGGTGGTTTCCCCCATCAGGAACCTGC
CAGGTACTGGGCCTGGATGGGCTTGAATTCCCCCATCTCAGTGGGAGAAAGGGGGCATTTGTT
TAG

SEQ.ID.NO. 10 HG1014912N1 PLT00014330B04.contig.b
ATGGGTCCCTGCGTTGGCGGGGGGACCCCTCGTTTCGCTCTGGCCTAAGCAGCCCCCTGG

GGTCCCGCGCCGCAAGAGCGCCAGACTTCCTAACCTGGGCCATACCCCTGTCCGCTCCA
 CTTCCAGTCCAGAAAGGACAGCCGAGGCCAGGACGCCAAGGACCTGGCTCGAAGTCT
 CCAGAAATTCGCCAGAGCTGGGCTGGGCTCGCGGACCGCTCATGCGCCGTCTCCCCAGGA
 GCCTCCCCCAGCCCCAGTTTAATAGCCGGCCCTGACCAAGTCCAGGGGAGAGACCGTGC
 AGCCTCCAGCAGACCGTGACCTCTGA

SEQ.ID.NO. 11 HG1014913N1 PLT00014330B05.contig.a
 ATGAGGTCAAAACCTAATAACAGATATATTACACTCTAACAGATACATTACATTGTATAGT
 AGAATTCTCAGACAGCAACCTTATAAGGTGATAGATGCAGAAACAGGCCCAAGAGTAAA
 AATAATTGTTACAGATTTTCCAGCCAAAGTGGCGAAGGTGAGATTCTTACCTGGGTTTGT
 CTGAAACCAATTTTCACTATCTTTTTCAGTTATGATACACTGCCTATTACGAATGAATGA

SEQ.ID.NO. 12 HG1014914N1 PLT00014330B11.contig.a
 ATGGCCCCCGCGGGCTCCGTCCCTCTTGCCGTCAGGGCCTCAAGCCCCCGCGGACA
 CCCACCTCGTACGGCGTCTTCTGCAAGGGGCTCTCCCGCACCCCTGCTCGCCTTCTTCGAG
 CTGGCCTGGCAGCTGCGCATGAACCTTCCCCTACTTCTACGTGCGGGCTCGGTGATCCTC
 AACATCCGATTGCAGGTACATAATTTAG

SEQ.ID.NO. 13 HG1014915N1 PLT00014330B13.contig.a
 ATGTCTTCCCCTTTTACTCATAGTCAGGACCCAGAAACTTCAAACGGCCTACAAGACTCTA
 CACCAGTGGCCCCCAACCTTTTGGCACCAAGGACCAAGTTTGTGGAAGACAATTTTCCC

ACTGACTGGGAGGTGGGGTGGGGCTGTGTGGGGGTGTGAGATGGT

SEQ.ID.NO. 14 HG1014916N1 PLT00014330B13.contig.b
ATGGAATTTTACAGGTCTGACATCTTGACAGAGGTATATTGTAAATAAGATACTCTCTC
CGGGAAAAGAGAAAGGCAGTTTAGAGGACAGGTAGAGAGGAAAGTATACAGATAAGGTGTGC
AGGTCTGCTCAGGGTTCCGAAGCAGTTTCCCTGGAAAAC TTGA

SEQ.ID.NO. 15 HG1014917N1 PLT00014330B18.contig.a
ATGCTGTTTTTCCCTTCTCATTCATCTTGACATTGTCTATTTTGAGATCTCAGCTCAGT
TGCAGAGAAAGCTGCCTTCCCCATCTGGCAACCCATTATATGTGTACATAGTACCATGTTGT
ACCATAGAGCTAGACACAGGTGCCATGTTGTGCTTGAATGTAATTCACCAGCTTCCAAA
GGTTTACCTCAATCCCCCTTTACTCAAGAACCTCAGTTCTACTGA

SEQ.ID.NO. 16 HG1014918N1 PLT00014330B18.contig.b
ATGCTAGATGTAGGGATAAATGGTTAACAGGACACAGTTCTTGCTTATCTTGCTATGGT
CTGTCAAGCCTTATACAATCTAGCCTTATCTCTCACCGTCATTATCITTATCACCCCGTA
GATTCCCTTGTTGGCCACTGGTTTCTTTCAGTCCTTAATTAG

SEQ.ID.NO. 17 HG1014919N1 PLT00014330C06.contig.a
ATGTTTGTGTCTGTGGTATGTTTTGTTACTGTGGCCCAAGCTGTTGGACACAGGGACAC
CCCTGGGCTCCCAGGCACCACCTTCCCTCACGGTGTCTCTCCTGGAAACACGGTCTGTG
CCTCACTGCCCATCCAAAGCCACATCCCGGCTCCACCTGCTCACTCCTGTGTTTCTTCCAC

TCGATGGTCAAATTTTCTCTTACCAAGAGCTGGCCAGGTGCAGTGGCTCACACCTGTA
ATCCAGCACTTTGGGAGGCTGAACAGGCAGATCATTTGAGGTCAGGAGTTCGAGACCA
GCCTGA

SEQ.ID.NO. 18 HG1014920N1 PLT00014330C06.contig.b
ATGACAGTTTCAAGAGCACCCACGTACTTGACCCCGCTGGGCAGAAAGTACCAAAGGA
GACCCCGGAGGGTCAAAGGAGGGTGTCTCAGGCCCGCGCTGAGCCCTGCAGGATGCTC
AGGCCAGGGCCCCACCCCTCTACTGGCCCTTCTCCCTCAGAGTTGTCCATTGACTGGACA
AACCATCCCCCAATGCTGA

SEQ.ID.NO. 19 HG1014921N1 PLT00014330C12.contig.a
ATGAGCCTTTCAGAAAGGCCCCAGCCCTGGGGATCTACGACCCCTGGTTCCCTTCAAGA
AAGCTCAGACAGCGAAGGAAGAAAGGGCAGGGCAGCCTATGAGAGAGAGAAGGCAGAGA
AAAAGGAAAGGAGATGGGAGGAAAGGGAAACGCCCTTCTTAGAACTCAGGAAAAACAACAT
CAGAAAGTGGGCATTTCTCCCTGTATAA

SEQ.ID.NO. 20 HG1014922N1 PLT00014330C14.contig.a
ATGCAATTATCAAACAACCCAGTCAACCAGGATTGAAAGTGTCAAACAGTGCAACAATC
AAGCCATTCTTTGTATGATCAAGAGCACAGAGCCTGGAGTCAGACTGACAAAGGTTTG
AGATTGATGAAAGATTGTCAGGTGACTTTTGGTCTCTCCTACATAGTCTGGAAATCAGA
ATTGGAGCTTTAAACTTTAA

SEQ.ID.NO. 21 HG1014923N1 PLT00014330C18.contig.a
 ATGCTGGGCAGCGTGGTCCAGCCCCCTGCCTCGGGGTGGAAGGCAGAGAGCAAAGAGGA
 GGCCCTCTTATTCCGGACCCCCAGGCCAGAGCCCAGGAGTGGAAAGCAGAGGGCAAAG
 AGGAGGCCCCCTCTCATCCCTGACCCCCCAGGCCAGAGCCTAAGGTTCAAGTGCCTCAG
 GCCCAGTCCCCCTTGA

SEQ.ID.NO. 22 HG1014924N1 PLT00014330C18.contig.b
 ATGAGCACACTGAGCTCTGCAGGGCCTCAGCTGGAAGGCGCCATCCAATCGTCCTTCCCC
 TCCCCAGCAAACTCCAGAGGTGGGGCGGTGCCGAGGGCCCCAAGGCTGCCCTCGGCAGG
 GGGCTTTGTGCTTTTGTGTGGAGGCCACCAAGGATTACAGGAAGATCCTGAATGGGTTG
 TCAATTGACTGA

SEQ.ID.NO. 23 HG1014925N1 PLT00014330D03.contig.a
 ATGACAAATACCCCGTCTCTGGTATCTTTGCACTTGCCCTCTTCAAAGCCACCTTTTCCCTT
 CTCTCTCTGGACTGTCACAGATCCTCCTCTGTACCCCAAGAACACCCCAATTGGATTTC
 CGGGCTGGCTTGGCCTACACACCTGGGTACGGCCTCTCGCAGGGATGCGCCTGCCAC
 TGTAATAAGAGGAGAAACGTACACGGGAAAGGCCTGACTCCTTCGTACAATCAGAAT
 ACTCAACTGCACAGAGAGGGACCCCTTTAAGCCACTTTGGGAGCCACATCCACCACTCT
 GTGACTCCACACAGGCTGGTTCACAGGTATCAGGTGTCTAGTGTAAACACGGGCCAAA
 AGGACAAACAGGACCCTAGGTCTTCCCCAAACATTAG

SEQ.ID.NO. 24 HG1014926N1 PLT00014330D03.contig.b

ATGTGCATTTCTCTGCTGAGATTCCACATGTATTTCATTCATTATGAGAAATTTTCTT
 CACCCCATGACTATAGTTTTTAATAGCTGCTCTCAAATACTTGACTGCTGATAACAACATC
 TTGGACATTTTGGGGATAGCTTTTAATGCCTATGTTTTATCTTGTGTATGGATAACATTT
 TTGTGTTTTCTCTCATGTCTCTTAAATTTTAAATTTGATATTTTAA

SEQ.ID.NO. 25 HG1014927N1 PLT00014330D05.contig.a
 ATGAATACCAGGAGTGGAACCACTAGAGACTGGCTACCATGGTCACCTACACCAAC
 ATGAAACTTTCAGCCTTTCCACCTCCTACCACTTATACTATATTTAATGTTTTCTCTGT
 ATGGACTTTAAAAAGAGGTCATCATAAGCAATAATATCCATGAAATCATGGATATCATG
 TGA

SEQ.ID.NO. 26 HG1014928N1 PLT00014330D05.contig.b
 ATGGGATCCCTATCCTTGTTTCCAGTTTCGGTCATGGTATTGCTTATCCATTATTAGT
 TATCCACTGCTGGGTACAAATTACTCCCAAACCTTAGCAGCTTCAAACAACAATAAACAT
 GTATTATGTCACATAGTTTGTTCAGAACTTTGGGAGTGGCTTAG

SEQ.ID.NO. 27 HG1014929N1 PLT00014330D07.contig.a
 ATGGTTTGTGTTATACCTGATCTGTTTTCTCTGTATCAAAAATAAAATTAGGATGAATCTA
 AAAGTTTATTGTCATAGAAAAGTACAGATCAGAAATCCAGGAGACTTCAAATCAAGTGGG
 AAGAGTGGTAAGAAAGCTTGCACTCACCAAGTAACAGCACAGCTTATAAGGATAAAGGAGGA
 AGCATTTTGAGCTTTTCCATTATTGGTCCTTATACATTAAAGATTTTGTGGGGGGTGGCA
 AACAGCTGTTTAAGTTGA

SEQ.ID.NO. 28 HG1014930N1 PLT00014330D10.contig.a
 ATGTTTTAGTAGAGATGGGTTTCATCATGTTGGCCAGGCTGGTCTCCAACTCCTGACC
 TCAAGCCATACACTGGCCTCGGCCCTCCCTAAGTGCTGGGATTACAGGTACGAGCCACCGT
 GCCTGGCCAGGAATCTGCAATTTTGTAGTAACCCCTGGTCATCTCAGTGCACACCAAAGT
 TTGAAACCACTGTTGGGGTGTGGAGTTCTTGCTTCCAGACTTCCAGGGCCTATGGATTGA

SEQ.ID.NO. 29 HG1014931N1 PLT00014330D10.contig.b
 ATGAACCTACTGCTGACCTTGGTTTATTTCGTTTGTATTATTTTCACTTTTCTGTGAAA
 TGTGCAAAACATACTCAAAAGTCAAGAGAATCATATGGTGATTTCATCATAGTCCCTTTGG
 CCAGCTTCAGCAAAAGATGACATTTTGCCACTCTTATCCATCCATTTCTGTACGCCACC
 CCTGACATACACACATGTTCTTCTGGAGTATTTTCAAGC

SEQ.ID.NO. 30 HG1014932N1 PLT00014330D12.contig.a
 ATGAGGAAGATGACATTTTGTCTGCAGATGGTGGAATAAAAAATCACAGAGATTGTGAT
 TTCCTTATGGTTTTGTGGTAATGGTGGAGTTTAAACTTACAACGAAGTTTCTGGTGACA
 TGTTTCCTAGTTTTCACAGAAAACATTCTTTTTTTTTTTGAGACAGAAAGTCTCACTGTG
 TCGCCCAGGCTGGAGTGCAGTGGTGTGATGTGTCGGCTCACTGCAACCTCTGCCCTCTGG
 ATTCAAGTATTCTCATGCGTCAGCCTCCAAGTAGCTGGGATTACAGGGGCCCAACCA
 CGCCCAGCTAATTTTGTATTTTAGTATAGATGGAGTTTCACCATGTTGA

SEQ.ID.NO. 31 HG1014933N1 PLT00014330D12.contig.b
 ATGGAGTTTCAACCATGTTGACCAAGGCTGGTCTTGAACTCCTGGCCTCAAGTGACCCACCT
 GTCTCAGCCTCCCAAAGTGTGGGATTACAGGCATAAGCCACCACCCAGCCAAAACA
 TTCCTTATAATGATACAAGTAATATACCAGACAAGAAATTAC

SEQ.ID.NO. 32 HG1014934N1 PLT00014330D13.contig.a
 ATGAAAGCTACAATGTCTCTTTTGAAACAAGTCGTTTATCAAGAGAAACATCTATCAGGGA
 CTGCCCTTGTCACATTGATCCTCTAAAGTTAGTCCGAGGTGGAAGCTGCTTTGATGTTT
 TTGATCCTAAATTTAATACAACAGAAAAAAGCAAGGATATTTTGACCAGTATAGATTTT
 TAA

SEQ.ID.NO. 33 HG1014935N1 PLT00014330D15.contig.a
 ATGGCTAATAAGCACATGAAAAAATGCTCACCATCTTAGTCATCAAGAAATGCAAATCA
 AAATATGAGATAGTATTATACTTCACACGCACTGGGATGGCTGTAAATCGAAAAGACAGAT
 GATAACAAGTATTGGAGAGGAGGTGAAGAAATTTAGAAACAGTCTGGTGTTCCTCCAAAAG
 ATAAACATAGAGTCATATATAACCTAACAAATTCCTAGATATATATCCAAGAGCA
 ATGAGAACATACGTTTCACACAAAAGTTTGACATTTTAA

SEQ.ID.NO. 34 HG1014936N1 PLT00014330D15.contig.b
 ATGTATGACCACGGGTAAAAAAGCTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGAG
 CCTTGCTTTGTCGCCATGCTGGGGTGCA GTGGCGCAATCTTGGCTCACTGCAACCTCCAC
 CTCCTGGGTTCAAGAGATTCTCCTGCCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCACA

CGCCACCATGCCCAGCTAATTTTGTATTTTAGTAGAGATGGGGTTTCACCATGTTGGC
CAGGATGGTCTTGATCTGTTGACCTCGTGA

SEQ.ID.NO. 35 HG1014937N1 PLT00014330D17.contig.a
ATGGGCTGCCCCGGTGAGCATAGCTGAGATGCACCAAGGCCACATAAGTCACCTACTTTGC
CTTGGCTGCCCCCATCTGCATCTATCAAAGGAAGCCTTGGACACCCACCAAGAGGAGCCAGC
ATGAGGGAGTGTATGCCTTTATACAAATTCACTCCAACCTCAGAAAAACGTCCGCAGCTC
ATGCTCCCCCTGCCAGAGCAGCAGTGTGAGCAGCTGTGTAGGTTTGGAAAGCACCCCAAGTC
ACTTGGGCATTGATATGGTTTGGCTGTATCCCCACCCAAATCTCATCTTGA

SEQ.ID.NO. 36 HG1014938N1 PLT00014330E04.contig.a
ATGCTGGATTTTACACACAGATCTGGTTTCAGGAAGAAACAGGATGCCAGTGCAGTAGCC
CTGTACACTATCACCTTAAAGGACCTCCGAAAGTTGTTTCTGCTTATCTCTCCATAACCAA
CTCCCTCCTTGCTGCACACTTCAGAAATTTGGAAATAAGACATAA

SEQ.ID.NO. 37 HG1014939N1 PLT00014330E14.contig.a
ATGATGCTTGAGACACACCAAGATACCTGCACACTAGAGGCTGCAGTCACCTCAGAAAGAA
AGATGCCCCCTGCTCTGGCTGTCAAACGGAACCAAGTCCGTCTTCCTGAGAGGTTTGGT
CCCCTTCAACCAGCTACAGCAGGGCTGGCAATGCCCAAGTCCTTGGAGAAACAGAAAGAGAT
TCAACTGCAACTGAAATTACCTACTAA

SEQ.ID.NO. 38 HG1014940N1 PLT00014330E14.contig.b
 ATGGAACATAAGAAATGAAGGAGAAATGGGTGGACAGGGCACAAATTGGGGAATGACAGCC
 TATTCAGTGAGGATAGGAAATGCACATAAAGGACAGATGAAGATTTCTCTGTGTGTG
 ATTAGAAGGATAGCAATTACAAAGATAAGAGTGAAAGGAGAGGGAGAAAGGGGAAAGG
 GAATAA

SEQ.ID.NO. 39 HG1014941N1 PLT00014330E24.contig.a
 ATGCATTTTGGGGACACATTCAGCCCATAAACAAGTACCACAATCTACCTTGGTTCTTC
 GACTTTACCATGGAGAAATTTATTTTCCCTTCCCCAGGACAGGCAGGATTGGAAAGACCTTT
 CCTGGTGCAGCCTGGAGAGCCCTGCTCACAGCCCACAGACAGGCTGCCTGCCTCACAC
 CCTACTCAGGCCAGGGAGCTGACCCCAAGAATCACAGGGCTGGGCACAGCAGCCTCCCTGC
 TCCCCACTGGCCCCCTGCTGTGAGTCACTCACAGGTCTCTAGAGGACAAGGTTCTCACTTT
 TCAAAATTATCTTTCTTACATATACACTTTTACCCAAATGA

SEQ.ID.NO. 40 HG1014942N1 PLT00014330E24.contig.b
 ATGTATACATATATATGTGTGTATAAGTGACTATAAGTGATACATACATGTACATAAGA
 TTTTCACATTTATTGTTATCTTTTTATTGTCAATGAAATGTCTCAGAAAAGAAATAAAAAAG
 AAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA
 AGAAAAA

SEQ.ID.NO. 41 HG1014943N1 PLT00014330F01.contig.a
 ATGGAATCTTGGATTGCCCTCCTTTTGGCCTATGAACATGGATATATATTTTCTACCAATT

TTTCATGAGTTAAAAATCAGGCTTGTAATTTACACATTTTTTTTCTAATTTACTTTTGG
ATCTTTTGTATTTC AAGGACCTAGGGTATGGTCTGATAGATAGAATTGACAGTGGGA
ATGAATCAGAGGCAAGGAGCATATAAAGTCAGTAGAGCCACCACATACATTGCTAA

SEQ.ID.NO. 42 HG1014944N1 PLT00014330F03.contig.a
ATGCAGCAGAGGCAAAACCATGGTTTGGCCAAAATGGAACAGAAATTTACAAAGGAAAAGG
GAA GTTCCTGGAGTAATATGTGTAGGTCAGATGAGGATGGAACTGGTGAATGAGCAGAAAG
AGAAAGGTTGGAAGAGTGCAGTGGTGATGACAGCCCTCCGCTGCTCCTTCCCCCAGAGGAAG
GAGCTCCAAAGAGATTCTCAACTGGCTGAAGGATACACAAGATGTCTTATTTGTGTGTAC
TTGTTTGTGTACTAATGACTCCTGTGTACACCACCTTGCTTCTTCAGGCAACAAATGC
TGTTTGCAGCACTTGTGA

SEQ.ID.NO. 43 HG1014945N1 PLT00014330F03.contig.b
ATGGAAGGTGACTCCTTTGAGCTTTGTGTCAGGTCAGCCCATAGCCTATGCAGTGTGCAAG
GTGACACCTTCCCTTAGGGGCTAGCGTCCCTTTCTGCTCCTTTCCAGCCTGGACGTCCCC
TGCTGTTGCGCTTCGGTCTTTTTTCGGCCTCCTCCCTGATGTGGCTTTGTCTGTCTCTGC
TTAATGCCACAGGCTTTGGTTCCTGCCCACCAATGA

SEQ.ID.NO. 44 HG1014946N1 PLT00014330F04.contig.a
ATGAATATGCTGTACATTGCTTTGCTGTTGTTACCCCCCTTAACCATGCTAGAGAATCA
CCCACAGAAAGGGTGTCTCCACCGTACACATACTACTTGGTCAGGGAACAGCATAACTAAA
ACTCTATTGTACCATACTTATTATGGGTGTATGGGAATCGCTTAGGAACTTGTACTTAT

AGTCAAAACCACTACTCAGTTTGTGACCCAGGAAATAACCAACTTTATGTATGTTATGAC
 CCCAAGTTTTCACCTGGTGAAATGTTTGAAATTTCGTGCAGTCAAAAGAAAGGTCCTCCTCTT
 AAACCAAAACCAAGGTCCCTCCCTTTTACTGAGGGACTATTCTCTGTATTTTGA

SEQ.ID.NO. 45 HG1014947N1 PLT00014330F04.contig.b
 ATGAGGAACGCTATCTATCAGAACAGATTAGCTTTAGACTATCTCCTAGCCCTAGAAGGA
 GTAGTATGTGGAAAGTTCAATTTAACAAATTGCTGCCCTGGAAATCGATGACAAATGGAAAG
 GCAATTATGGAAATAACTGCAAGAAATGAGAAATTAGCCCATGTTCCAGTTCAAAACTTGG
 AAAGGTGGTCTCCAGATTCTCTCTTTGGAGGCTTGGTTTTTATTTTCTGAGGGTTCAA
 GACTTTAATAGGAGTGGTTCTGGCCATAATTAGGAAGTTGCCTAATACTCCCTTGTCTCTT
 ACCTCTCCTTGTTAG

SEQ.ID.NO. 46 HG1014948N1 PLT00014330F05.contig.a
 ATGTCTGTAAATCCAAACACTTTGGGAGGCCGATGCGGGCAGATCACTTGAGGTCAGGAGT
 TCAAGACCAGCTCGACCAACATGGCAAAACCCCGTCTCTACTAAAAATATGAAAAAGAAA
 AAATTAGGCTGGTGTGCTGGCACATGCTGTAG

SEQ.ID.NO. 47 HG1014949N1 PLT00014330F13.contig.a
 ATGGGAGGTACGCTCCTGTCCCTCTCGCCCGCGTCGGCTAGTGTGCTAGTCTGGCGCGG
 GCAGTGTACCGGTTGCTGGGAAGACTGGGAGGCCAGCGAGCGGTACCAACCACTCAG
 CCTCTCCGCCCCAGCTGGAGCCGGGAAGCAGCGGGGCTAA

SEQ.ID.NO. 48 HG1014950N1 PLT00014330G21.contig.a
 ATGGTTCCACATCACCTTTCTTCGAGGGACAGACTCCTCTTCTTGACCTCCGTCTCCATG
 TTCCCTCATGACACTTCACTGTCTTCTCTCCAGCATGGACACTCTTTCCATCACTACCGCC
 CTGGAGAAAGCTAGTCCCTCTTCTGGGAAGTCACCTCCTTCCCAAAACCTCACACGAATTTT
 CTGAGGAGCAGCCCCAGTGCCCCCTTCTCAGCCAACACACCCACCTCCATTCTCTCTGAG
 CATCCAGTTTGTGTCAGCACAGTTCTGTCTCTCCCCCTTCCAAGCTGTTCCCAAGGTTATC
 ATCCAGGGATATGTAAGTTTCCCATCTGAAATGCATCATTCCTCCTCCTAGGCAGTCAA
 ATCTCATTTCCAGGCCCACATCATCTGATTAATTTCTTAAACATCAGAAATATCATTAAC
 AATGATACATCAGTTTCTTAG

SEQ.ID.NO. 49 HG1014951N1 PLT00014330G21.contig.b
 ATGTATATATGTATATATTTATGTATATACAAATGATGAAAAACCGACATTTTACACTA
 TATACATATGTATGTATATCTATGGACAAAACCAATATTTTCTATGAATTGTTTATTC
 CTATTAATCAGCTATAGAATATGTATGTTACATAATATATAA

SEQ.ID.NO. 50 HG1014952N1 PLT00014330H05.contig.b
 ATGAATAGCAGGTTTGAAATCAGTAAACAAAATGTCTCCCAAAGGAGAAAAGCCCTAGACTA
 GGCTTTTATGCTGATTTCTACCCCAATTTATAAGAAAAACAAACACCAATTTCTTCTCAA
 CTATTCCCAAAAATTGAAGAGGAAGGAATTTCTTCTAACTCATTTGTATAAGGCCAGCATT
 ACCCTGATATCCAATCAAGACAAGGACACAAACAAAAGGAGAAAACTACAGGCCAAATATTC
 CTAATGAACACAGATGGTAAAATTCTCAGCATAATACTACCAAGCCAAATCTAA

SEQ.ID.NO. 51 HG1014953N1 PLT00014330H06.contig.a
 ATGAATATAGAAATGGAAATTGGTCCCTTTTCACCGATCACAAACTTTGTACCTCAATTC
 TTAAGTCAGGGATGTTTAACTCAGCCAGCCACCTAGCTAGGGTCTGTAAACAAACAGTTG
 GTTATTTTTCACCCCTTTGTCTACTCAAAATTAA

SEQ.ID.NO. 52 HG1014954N1 PLT00014330H12.contig.a
 ATGGGGCTGCTCCCGCAGGCTGGTCTCGCCGGTCTGGGGGGCTGGGGCTGGCCAGG
 CCGTCTCTTAAAGCGCCGCGGGCGCGTCCGGCGAGCACAGGGCGGTCCGTGGTGCTG
 CTGGTGCCCACTGCGGCGGGGAGCCGCGCAGTGCAAGGCAGGGGCAAGGACCCAGCG
 CCCCAGGCCGAGTCCC GCCCCAGCCCCGTCCCTGGCCCCAGGCCAGGCCCGCGGTGGA
 GGGGCGCAGAGCCGACCGTGA

SEQ.ID.NO. 53 HG1014955N1 PLT00014330H12.contig.b
 ATCGGGGCTCGCGCACCCGTTTCGCTCTCAGCGGACAAACCGGTCGAGAGCGCAATTCTG
 GAGCGGGTTCCAGAAAGCTCCTGGGTTCCAATTCAAGGCTCCTCACTCGCCAGCCCTGCGAT
 TTGGGACAACTTAACGTTTCGTAAGCTGGGTTCACTCTCTCAATTGGGTTAATGTCCGTG
 AGGGGCTCTCGTAAGGGCGGCTGGCCTGCCGTAACCTTAAGAAGCATGTAA

SEQ.ID.NO. 54 HG1014956N1 PLT00014330H14.contig.a
 ATGCTGGTCATCATAGGCTCTCCTGTGCAAAACAGCCTTGTATAGTGGGTTGAGAGATTAC
 TGTCTCAGTTACCCACTCCTGCTTGGAACCAAGGATTCTGTCACTGTCCCCAGAGGTTGTG
 GATCAAAATCAATGTGCCTGTTAGACACAAGCCCAAGAGAAGGAAAGGTGTGCTTGAATTCA

TTCACGATCATGTCATGTAACGTTTAA

SEQ.ID.NO. 55 HG1014957N1 PLT00014330H14.contig.b
 ATGAGCATGTGCGTGTGTTACAACTTTTCATTAAATAAATCAACTCCTGGGTCAATTGTCC
 ACTTTTAAACTTTTCATTGTGCTCATAATTTCAGATTTACAGAAGAGTTGCAAAATAGT
 TCCTATATGCCCTTTCACCCAGATTCCACTAGTGTAGTATTTTACATGACCATGGTATAT
 TTATCAAAACTAAGAAAATAA

SEQ.ID.NO. 56 HG1014958N1 PLT00014330H18.contig.a
 ATGGCCCTCCTTTGTACATTTTCTTTTACTGTTCTCTTACCCCTTCTCACTCTCACT
 TCACTTCCTCCATGCTGTCTACTACCAGTAGCTCCTCTTACCAAGAGGTTCTATGGAGA
 ATGTGGCTTCCCAGAAATATTGATGTCCCATCGTATAGGGGTTTTTCTAAAGGAGACCCC
 ACTTTCACCCACCAACCATATACCCCTGCACCTCAGGCCATACATTTCAATCCCTGTA
 TCTTTAACCTCCTTGTAAAGTATGTCTCTTCCAGAATCGAAGTTGTAA

SEQ.ID.NO. 57 HG1014959N1 PLT00014330H18.contig.b
 ATGAGAAATTAAGGAATAATCGCCATAAAGGCCTGATACTGGCAATAATATATTGAC
 AAATATGAGTTTCTCTTTTCCCTTTATTCTACTTTACGGCAACTCCTACCTCCTATTCT
 CTTTCTTTTGCTACTAAGCAGACATCTGGGTTTGTGTTTGTGTTTCATCAGGGAGAGGTT
 CTAAACGTTGTATCCTTTAACCCCTCAAAGCATCATCTATTCTCTTAGAGGGTAA

SEQ.ID.NO. 58 HG1014960N1 PLT00014330H11.contig.a

CGCGCCAGCCGAACGGGTACCGACCGTACCCCGGATCTTCCTCGCAGGGGGCGCCG
 GCGCCCAAGCTGCTCGCCCGAGGGCCAGAGCCGCTTCCAGAACAGAGCCGGGGAGG
 CGCAGCCAGTAG

SEQ.ID.NO. 62 HG1014964N1 PLT00014330I13.contig.b
 ATGCATTCAGATCTCATGTGCATTTTCTCGTTATTCAAGAGAAATTGGACACTTGATGAAG
 GGAAGGACACGGAGAGATGACAAACGCAGACTGGCATGCAGTTGTTTTCAGGGTGGGAAG
 AACAAAGTGGAATGTCTTCAGTTCTTACAAATAATGTTGGCATTTGTAAATCTCTGAGTTT
 TTGTATTTGGCCTACAAAGGGGGTCTTTTCAGGGGAAACACGAATGGAAACCTTCAGAGAC
 ATGTCTACATTGTAA

SEQ.ID.NO. 63 HG1014965N1 PLT00014330J10.contig.a
 ATGGCCTTCTCCAATTCTTCTCTCCAAGCCCATGGCAAGGCCAATGGGAAAAAGC
 GAAACCTTAAAGGCGAGCCTATCATTTATCCCAAAGCCCAAAGAACAAAGGCTATGTGG
 GTACAGTGGATTTGCTCCAGAAAGCCCGGCCTCTGTCCTCGTGTGTTCTGAAAACCATCCT
 TTCTGGCAAGGAAAGAAAGTCTCAGCGCTTTGGATGAAGCTCTCTGTTAGGCCACCACTG
 CTCTCACGCCCTTCGACTGCTGGTGCAGGAAGATTCCAGGCCACTGGGCTTCACCGGCC
 TCAAATGCTCACTTGTCCAGCCACAGCTTGACAAACTGGTCACTTCCACCCCACTGCTG
 GAGCTGACTAGGCAAGTAAGGCTGATGGGTGA

SEQ.ID.NO. 64 HG1014966N1 PLT00014330J10.contig.b
 ATGGGTGACCCCTGATCGGCCCTCACTGGAGTACAGGGTTTCCTTCTCCAATCTCCTCAA

ACAGGTGAACAAAGTTCTCAGGCCAAGTTGCTCCAGCACCTCCCCAGTCTCAGACAAAGT
CCATGGACCATCTCTCCTGCTCCATGCAGGAGGATTCTCCAACCAACAGTGCCAGTTCTAC
AGGTGTTATCACCCAGCTAGGAAGGATTTTGAGAAAGAGCAGCTCTGTCCTATGGGCTGTG
GGTTCACATCTCCACTCTCCACCCAGACAGCCCAAGGAAGGCTGTCTCCCTTTGTTCGG
GCTGAGGTCTGA

SEQ.ID.NO. 65 HG1014967N1 PLT00014330J14.contig.a

ATGCTGTTTTCCATAGTGGGTGAACCATTTTACATTCCACAGCAGGTGCACCATTTTATA
TTGCATAGTGGGTGCACCAATTTTAGTTCTCTGCCAACAGTGCACAGTGGCTTCAATATCTC
CTCCTCCTCACCAACACTTGCTACCTTCTGTTTCTTGATAATGGCCATCTTAGTAGGTGT
GAGGCTTATCTTACACCATACACAAAAATAAACTCAAAATGGATTAAAGGCGTAGCATAA

SEQ.ID.NO. 66 HG1014968N1 PLT00014330J14.contig.b

ATGTTTGCAACATGGCATAAGTATTGCTAATTACATACATGGGGCTGTACAGCACGCCTCC
AGGGCTCTTCCATCTCACGTGGGTGAAACTCTCCACGTTCTCTCCCCCAACCCCTGGCAAC
CACTATTCTAGTCTCTGTCTATAAGTTTGACTGTTTTAGAGACCTCATATAG

SEQ.ID.NO. 67 HG1014969N1 PLT00014330J15.contig.a

ATGAAGGAAATCCATAGAAAGGTCTCTTAGTGAAACAAAATTTAGTTATTAACTTTATA
GCTATGAAATTTCCCCGGGCATTTGTTTTTGTTCAAAACAGACTTTAACCTCTGCATCATA
CTTAACCCCTGCGACATGCGTACAGTATGCATATTTTGTGTTTGAAAAAAATGTTTCGTTTC

CAGTCTGTTAAGAAATATTCAAAAATAA

SEQ.ID.NO. 68 HG1014970N1 PLT00014330J21.contig.a
ATGGGCAGTATTAGAAAATAATTTAAATGTATATACCCCTTTGACTTACTGGTTTCCTTCT
ACCACCTTATTCTGTAGATACACCCACATCTACATTATTACAGCATTTGTTTATGGCAGC
AAAAAATTGGAAACAATGGCAGGTACCACCAAAAAAGAGACTGGTTGAATAAACAAAAAG
CATAGCTATGCAGTGGAATGCTATGCAGCTGGTAGAAAGAGGGGCAGATTGATATGTAC
TAA

SEQ.ID.NO. 69 HG1014971N1 PLT00014330J21.contig.b
ATGGAGAAATGGCTGGGTTTGGTGGCCCATGCCTGTAAATCCCAGCACTTTGGGAGGCCGAG
GTGGGCAGATCACCTGAGGCCAGGAGTTCAAAGACCAGCCTGGCCAAACATGGCAAAACTCC
ATCTCTACTAAAAATACAAAGCCTAGCATGGTGGAACATGCCTGTAGTCCCAGCTACTTG
GGAGGGTGCGGCACGAGAAATCACTTGA

SEQ.ID.NO. 70 HG1014972N1 PLT00014330K01.contig.a
ATGGCACTTGCTTTCTTCCACATGTCTCCTACCAATTTTACAGAGCTCAACCCCAACCCCC
CACACTCCCACCCCCCAGGGATGGAGGCTCGCAGAGATTTCATCCATCGTTCAGAAA
AAGTCAGGAGAGAACAAAGAGGGGAGGGGATATTGCTGCTGTTTGGTGTATTTTG
TGTGTGTGTTTTTTTGTTCCTAGGAAAAAGAAATAGGCTCTAA

SEQ.ID.NO. 71 HG1014973N1 PLT00014330K08.contig.a
 ATGTGCCACCATGTCTTGCTAGTTTAAATTTTGTAGGAACAGGATCTCACTACAT
 TGCCCAGGCTGTCTCCAAAGCCTTGAGCAATCCTCCTGCCTCAGCCTCCCAGAGTGGTGG
 GATTACAGGTGTGAGCCACACACCTGGCCTTTTGTGCTTACCTTTCAATTCTGTTTGC
 TGTTTTTTGACATATAGGGGTTTTAAATTTGTGTCAGTTGGATTACATTGCCCTTTTGT
 ACTTTTTTCATTGCTTCTATAAAGCTAAATCTTCTTTTCCCATTCCAAATTCAGAAATAA

SEQ.ID.NO. 72 HG1014974N1 PLT00014330K08.contig.b
 ATGAAATCGAGCATTTTTTTCATATTAATTGGCCTTTATATTTTGTGTAATTGTCT
 GTTATTACATTTTACCTTTTACATTCAATACTATACAGAAATATTATCCTGTTACAGCTA
 TTTTATTTTGTGCTTACCTTTTATTATTAG

SEQ.ID.NO. 73 HG1014975N1 PLT00014330K09.contig.a
 ATGGCTAAAGGGGCCAAGGTACAGTTCTGGCCATTGCTCCGAGGGTGCAAGCTCCAAG
 CCTTGGCAGCTTCCACATGGTGTGGTTCTGTGAGTATGCAGAAAGACAACTGAGGTT
 CAGGAAACTTTGGCTGGATGTATGAAAATGCCCTGGATGTTTCAGACAGAAAGTTTACTGCAG
 GGATGGAGCCCCTTGTGGGAGAACCTCTGTCTAGGGCAGAGCAGAAAGGGAATGTGGGGTTG
 AAGTCCCCGCACAGAGTCCCCACTGTTGCACTGACCACTGGAGCAAGAAAAAGAGGGCTA
 TAG

SEQ.ID.NO. 74 HG1014976N1 PLT00014330K09.contig.b

ATGGGGCGCTACTATAAGGATACCAAAACATGGAAGCAACTTTGGAACTGGGTAATAGG
CAGAGGTTGAACAGTTTGAAGGCTCAGAAAGACAGGAAATGGGAAAGTTTGGAA
CTTCCTAGAAACTTGGAAGATTTCAGAAAGACAGGGAGATGTGGGAAAGTTTGGAACTTCCC
TAG

SEQ.ID.NO. 75 HG1014977N1 PLT00014330K15.contig.a
ATGATGCCGGGGAGACCCATTTCGGGGCGCCCGGACGGCGGGACCTGTCCGGATGT
CAGGGCTGCGCCTCTCTGCAGCAGAAATTTGAATGAATATGTTGAAGCATTAATTACCTTG
AAACAAAATAATTATCAATACAGATAATTTGTTAACAGAAATATCAGAAAGAAATGTGATGAT
ATCCTTTTACTGGCTTTTCTGTGAGTTTGTGAACTAA

SEQ.ID.NO. 76 HG1014978N1 PLT00014330K15.contig.b
ATGGCTGGAGGCCCCAGGAAACTTATAATCATGGCAGAGGGGAAGCAACACATCCTT
CACATGGTGGCAGGAGAGAGAAAGTGCAGAGCGAAGCGGGGAAAAGCCCCCTTATAAACCA
TCAGATCTCGTGGAACTCACTCATTAATCAGGAGAACAGCACTGGGGGACTGCCCCCGTG
ATCTGA

SEQ.ID.NO. 77 HG1014979N1 PLT00014330K24.contig.a
ATGTCAGTAAAGGGTGCTAAAATGTTAAACAAATATTACTTTCAGATTTTCTTCTCT
CTTTTTATTCCCTATTCTATTTTCCCCCTTGGTTAATTTTATTCTCAAAATTCTCCCT
TTCTTTTCTAAACCTTATATTCCCCCAAGAAATAG

SEQ.ID.NO. 78 HG1014980N1 PLT00014330L01.contig.a
 ATGCCCTGTTTTAAGATTTTTTTTTTCCACACAGTTTTTGCAATTAAGTACCAGAAATACCC
 AAACATTTTTCATTAGTCTTCATCAACACTCCCCTTGTGTCTCAGTCAAAATTAATAAT
 AGCAATTTTGGCCGGGTGCGGTGGCTCACACCTGTATCCCAGCGCTTTGGGAGGCCGA
 GGCATTTCCACTAAACTACAAAAAATTAGCCGGGTGTGGTGGCGGCTCCTGTGGTCC
 CAGCTGCCCTGGAGGCTGAGGAGGAGAAATGGTGTGAACCCGGGAGCGGAGCTTGCAGT
 GAGCTGAGGTCGCGCCACTGCACTCCAACCTGGGCAA

SEQ.ID.NO. 79 HG1015004N1 PLT00014330L24.contig.a
 ATGCAGGTTGTGGCAGAGGCAGATGGACCTGCACTGTGACTTGGCCAAACGTGGACTGT
 GCCGATTGTCTCCTTCTTAGAGCCTCAGTTTCCCTCCTCAGGAATACAGAGCAGCTGCCCTC
 CGTGACTGCTGATTAGAGGAGATGGTAGTTGA

SEQ.ID.NO. 80 HG1014981N1 PLT00014330M02.contig.a
 ATGGTGAGTTTCTATCCACTAGAGAGTCATAGCTGTCTAAAGGGACACCATGCACCATTT
 TTCCAGCCTTACCTGGTCCCTCCAGATGCCCCCGCAGATACTTGTGGTGTGGCCTTCGCTC
 TTGCAGGGTCTCAGCCCCCTGGGACATCCTCCTACATCCACATCTTCATGGCTGAATCCC
 ACCCAGACCTCAAAATGAAACGAAAGGCCACCTTCAGGAAGCCTTCCTCAATCCAGTGC
 TTCCCTTTTGTCTCCTCATCTTCTTCCCCATTAAAGGGCTTAAGAATGATGTCCTCCAAA
 TTTGACCACCTTGGCTTTTAA

SEQ.ID.NO. 81 HG1014982N1 PLT00014330M02.contig.b
 ATGGAAGGCTCCCCCATCCTGTCTCTACACCACCTTGACAGCTGGGCTCAGCAACTGG
 GCTTTAATTTCAGTCTAATTCAAGTCAGCAGCATAGGGCAGCTCCTGGGAAATTGGTTT
 ACACATGCGGACAAAGCCAGTAGCCAGAGCTAACCCACTCACCAATCCCTGACACACAGAG
 GAGCAGATAAGGAAGCAAGATTGTGCGGAGCTATGACAAAGTTCCACTGTGGATGGCAAG
 GAAATGAGGGGGGCAGCAACCTGTAA

SEQ.ID.NO. 82 HG1014983N1 PLT00014330M08.contig.a
 ATGTATCCATGTGTTCTCATTCAGCACCCACTTATAGGTGAGAACATGCAGTGTTTG
 GTTTTCGTGTTAGTTGCTGAGGATAATGGCTTCCAGCTCCATGCCCTGCAAAAG
 GACATGATCTCATTCCTTTTGTGGCTACACAGTATTCATGGTGTATATGTACCAAGTT
 TTCTTTATTCAGTTTGTGATGGCCATTGGGGCTAA

SEQ.ID.NO. 83 HG1014984N1 PLT00014330M08.contig.b
 ATGCTTCCACAATGGTTGAACATAATTACATTTCCACCAACGGGTGTAAGTGCTCCTT
 TTCTTTCACAGCCTTGCCAGCATCTTTTGTCTTGACTTTTATTTTATTTATTTATTT
 ATTTATTTATTTTGCACTCCAGCCTGGGCAACACAGTGA

SEQ.ID.NO. 84 HG1014985N1 PLT00014330M15.contig.a
 ATGGAGAGAACAAAGGGAGCGGAGCGAGACCGGGCTGCGCCGCCCGCCGCTTGCG
 CACAGTCCCCGCCCGCGGGCGCGCGCGCCCTCCCTCCCGCGCCTCCTT
 CTGACCCCGAGCGAGAGACCTGGTCGGGAGTCCGGCGGCTAG

SEQ.ID.NO. 85 HG1014986N1 PLT00014330M17.contig.a
 ATGGCGCCGCGCGCTGCGGCCCGTGACCTGGACCACTGGGCTGGGACCCACGACCCCA
 CCACCACGCTGGAGGACGCGCAGCGGCCCTCGGCCCTTGGAAGGGAGATTTCACAGCGC
 GAGCCCCGCCCTTCCCCGCCGGAGGCGCCTCGGGCGGGGTCACTGGCCGGGAGGCC
 CGCCTTCCGCCCGAAAGCGCCGGTGACCCCGCAGGCCCTGCCCGTGGGTCCACAGGCCG
 GATCCCCGGGTTCCCTAGGCTGGGACTGCTTCTCAGAACCACTCTGTCGTTTTTAAGCAGG
 GTCACACACTCTAGCTCACTGGGTCCATTTAA

SEQ.ID.NO. 86 HG1014987N1 PLT00014330M17.contig.b
 ATGTTAGAGTCATCGAAAGTTTTTAATCAGGCAAGTGGTATGATCAGTTTTTTGTTTGT
 TTTGTTTTTTTGAGACGGAGTCTGGCTATATCGCCAGGCTGGAATGCAATGGCGCGATC
 TCTCCTCACTGCAACCTCCGCCCTCCGGGTTCAAGCGATTCTCCTGCCCTCAGCACCTCGA
 GTAGCTGGGACTACAGGCGCGGCCACCACGCCCAGATGAGAAAACCTGAGGCACAGAGAG
 GTGAAATAA

SEQ.ID.NO. 87 HG1014988N1 PLT00014330N10.contig.a
 ATGAAACCCAATCCCAGCAAAGAAAGTGAAGAGCAGATTATACAGTCCCATCCAAATTT
 CTCTTTGGCTTCTCTCTTTGGTCTTTTCATCTCTCTGCCCTTTCTCTCTGTGTCTCCTCT
 ACTCTTTCTTCTCTCTCATACACATACACACACACACACACACACACACGTATGTG
 CTTAGTGGTATGATTACTGTCAAGGAGAGGGCTACCTAGATGCTAA

SEQ.ID.NO. 88 HG1014989N1 PLT00014330N10.contig.b
 ATGCTAACACACACACACACACACACACACACACACACACACGCTTTGGGTGGT
 GGGAAAGTTGGAGAAAAACATTTGGGGCAAGAGAAAAACAATGAAGGCGGCCAGGTATGGTGG
 TTTGGCCCTGTAAATCCCGGCACATTTGGGAGGCTGAGGTGGGCAGATTGCTTGAGCTCAGG
 AGTTCAAGACCAAGCCTGGGCAACATAG

SEQ.ID.NO. 89 HG1014990N1 PLT00014330N12.contig.a
 ATGTCCCAGCAGCAACTGGGAACCTAATTCTGACCAACAAGGAGGAATTGTTGGCTA
 ATTGGAAGTGACAGAAACCTTGGGAGAAAGCAACCGCATACCAAGGATAATTATAAGAAG
 GTAACATGGGCCCTAAAGGATAGTGTCAGAAACCTGAAACCCCAATGA

SEQ.ID.NO. 90 HG1014991N1 PLT00014330N12.contig.b
 ATGATCAAGAAAGGAAATGCCCACTAGCAGCTTCTACTGTAAGGTTAGAGCTGACAGAG
 AAAGAACTCCTTAAGTACTAATTTGCTTCAATCTCTACAAAGGAAACTAGAGAAAGTGGGTT
 GACGTAATAGAAAGAACATGTGTTGTGGGCCAGGCAACCTGGGTTTAA

SEQ.ID.NO. 91 HG1014992N1 PLT00014330N13.contig.a
 ATGATGACTATGTTACCGGTTTTTGTGCTTACTATACTATACTGTTTACTGTTTAGAGTG
 TACTGTTCTTACTTATCAAAAAAATTAACTGCAAAACAGCTTCAGGCAGGTCCTTCAGGA
 GATGTTCCAGAAAGGCATTGCTATCATAGATGATGACAGCTCCATGCATGTTATTGCC
 CCAGAAAGAACTTTCAGCGGTACAAGATGTGGAGGTAGAAAGACAGTGATATTGATGATCCT
 GACCTTGTATAG

SEQ.ID.NO. 92	HG1014993N1	PLT00014330N13.contig.b
ATGGGAAGTGCTCTATATAAAAGTATAAATTTTTTTTATCTTGTATACCACATGTTCACT		
GTGCCCTTTTCCCATGTTTAGCAATGATTAGATACACAAATAGTTATCATCGTGGTACAACACT		
GCCTACAGTATTCAGTATAGTCACATGCTGCCGAGGTTTGTAGCCTAG		
SEQ.ID.NO. 93	HG1014994N1	PLT00014330N22.contig.a
ATGTCCGATGCTTTTTCAGACAGATTTTGGCTGCAGAGCTGGAATCACAAGATCATA		
ATGACAGCTCTAACAGCCGTGCCCTAGGCTGATGAAAAATGCCCTTTGAAACAGCCCGGGCTC		
ACAGGCATTATCTTTGTCTAATCCTATAATTGGCCTGACAAGTGGTGAAGCATAAATGAT		
TTCGGATATGTGCTTTTGTGTTTAAAGACTGGGATCTGGGGTGA		
SEQ.ID.NO. 94	HG1014995N1	PLT00014330N22.contig.b
ATGGGTCCCCAGTCATTCCCACCTGGTATTCAGACTAAAAGCAACATTGGTGGCAGGAAAG		
TCCACAGGAAAGGCTCTATTTTCAGGAAGAAGCCAGAAACCCCTGGAGAGAAATCCTAGTTCT		
GCCCCTAACCTGCAGGATGGCCCTTAGATGGATCTCTTTTGTACCTTACTTTGA		
SEQ.ID.NO. 95	HG1014996N1	PLT00014330O03.contig.a
ATGTGGGGCGAAGTTAGAGTAATAATTTTAGGTTTTATGGCTGGCTTAGGGGAAAAGGGA		
TTTTTTGTTTTTATGACCCACCTCGGGGAAGAGGGATTCAATTTCTATGGCTTGCTTCAG		
GAGAGAAATGAGGGATGAGAGACAGAGGGTAGGAGAAAGGTCAGAGAGAGAAACTTTACTT		
CTGAGGTCTTCACCTTTGGGGTATAATTTTCTGA		

SEQ.ID.NO. 96 HG1014997N1 PLT00014330O07.contig.a
 ATGCTTCCAAACACGCCCATTTCCAGTGATGATGTTTGTGCCACAGGCCAAGTACT
 GAACATCATGACACTCATAGACACTCTTCACTGAAGACGTGGATAGACCCCTGGGGCCACC
 ATTCTCTCCTCTCTGGAGTGTGGAGGGGCCCTGACCCAGGGCAGGACCCCCATGTCTGGA
 AGCAGGGTTAGCCCTGGCTCATGGGGGACCCCTGCTGGGCTTAGCCATTGCCACCTAG

SEQ.ID.NO. 97 HG1014998N1 PLT00014330O07.contig.b
 ATGCACAAGAGGCAACAAAGGCAAGTAAAGAGGAGGCTGGTGGGCTTGCAGAAACGCCCT
 GCAGTACCTATGCTGCGGTTCTCCCACTTCACTGATAACAAATCCTCAGTATCCACTGAT
 ACAAACTCTGGGACTGCTGGGCCAGTCCCAGGAATGACACTGGGACAGACTCCCTGGAGG
 TGGATCAGAAAGCTGCAGTTTACCAGGGGTCCTGCTGTGA

SEQ.ID.NO. 98 HG1015005N1 PLT00014330O18.contig.a
 ATGCCGGAAGAGATTGAGATTCCTAGAGAATGGAGGAGCCGGGAACAAGGAGAAAGGATG
 GGATCCAGGAACAGGCCAAGAGAACCAAGCTCGATGCTTCTTCTCTCCCTTCTAGAAAAT
 TCTGTTCAGAAAGCCATTAGGCTTCAGGCTCGAGACAGAGAGGAACCGCCTCTGTGTTTTT
 CACTGTCGTCTCAGTGTGGGAGTGTACTCCATGTGGTCAGTGTGTGATCAATTTAAC
 TGTGCCTGA

SEQ.ID.NO. 99 HG1015006N1 PLT00014330O18.contig.b
 ATGGAGGAGCCGGGAACAAAGGAGAAAGGATGGGATCCAGGAACAGGCCAAGAGAACCCAGC

TCGATGCTTCTTCTCCTTCTAGAAAATTCTGTTTCAGAAAGCCATTAGGCTTCAGGCT
CGAGACAGAGAGAAACCGCCTCTGTGTTTTTCACTGTCGTCTCAGTGTGGGAGTGTA
CTCCATGTGGTCAGTGTGTA

SEQ.ID.NO. 100 HG1014999N1 PLT00014330P07.contig.a
ATGTTGACCAAGCTGTCTCAAACTCCCAACCTCGTGATCTGCCTTCCTCGGCTCCCAA
AGTGTGGGATTGCAGGCGTGAGCCACTGCGCCCCGGCCTGTACCCCTCCTTCTTTTGAAA
AAGAGAGTACCTTCTTCTCTGCTTTTCATTCTGAAGCCTCCATTTAAACACTCCTCCTC
CTTAGCAATAGTGAGGGTCCCTTCAGACTAGATAGGATGTCAAAGGCCAAAGCATTTGCC
CCCAACACACTCCTTAA

SEQ.ID.NO. 101 HG1015000N1 PLT00014330P07.contig.b
ATGGTTGTTTGCTCAGAAAAGCCATCTCATCCCCCTCCCTGGCATCTTCTGATCTG
TCCTGTTCTGTAGTGGAATGTTCCCACTTGTGTTGCTGTTCAGACGTGTTGTCTT
CCATTATTTTCCGGCTGGATTGCCAGCAGAAATGCAGAGCCTGGAGAAATGAAACAG
AGCTGA

SEQ.ID.NO. 102 HG1015001N1 PLT00014330P09.contig.a
ATGAACACCTCAGCCATCCATGACAGCTGTTTTCTCCTCGGCTTTAAATCCTCCCTC
CTCAGCCTTCACCTCTACTCACCACCTCAACTTAATTCTCATTTCCCAACCCAGGAAGAA
CCTGTATGCCCCCTGCATCCCACTGCACTCCGGATGCCTTTTCCCCCAAGACTGGAATG
GCCTCAGGCCCCCATCTCCAGGTGAGTGGCCCTCACCTACAGACTCAACATATGGCCCTT

GGCTCTTCCCACTTCCAAGAGTCTTTGGAAGGGATGGGTCGAGCAAGCAGAGGAAAGGAAG
ATGTGA

SEQ.ID.NO. 103 HG1015002N1 PLT00014330P09.contig.b
ATGCCCCCTGCATCCCACCTGCACCTCCGGATGCCTTTTCCCCCAAGACTGGAATGGCCTC
AGCCCCCATCTCCAGGTGAGTGGCCCTCACCTACAGACTCAACATATGGCCTTTGGCTC
TTCCCACTTCCAAGAGTCTTGGAAAGGGATGGGTCGAGCAAGCAGAGGAAAGGAAGATGTG
AGTTCCCAAAATGCTCCTCACCTTTTCTCTGAGTGGGCTCCTTCTCAGTGGCATTGGA
GGGCTTGCGGCGCAGCATGGTCCCTCCACCCCTGGGAGACTCCGTCCTCCTCCTAG

SEQ.ID.NO. 104 HG1015003N1 PLT00014330P15.contig.a
ATGAGGTTGTGGGGGAATGTAAAACTTGAAGGACAAAGGTAGCTCAGTCCAACCAGATG
AGCATCCAAGGGAGATTTTGTCTCATCCTTTGCAGAAATCTCCATCTAGAGCCCGTGA
AACCAATGGAAATTCTTACTTGCTTTAACAGCACACTATTCCATTTCATTAAATGTG
GGCTAG

SEQ.ID.NO. 105 HG1015007N1 PLT00014333A03.contig.a
ATGTTTCGTGCTGCAGCAAAATTGGGGGGAACATTGGCAAAAATCTTCCAAGAACACATA
GGTCTTGGGATGGCCACATGCATTTGGCTATCTACATTCTCTGCACATCTGTGGAGAGAAA
AAAAGTTTCTTTTTTAAATACGGAGTCATACGTTTATTCTTAGGCTACTCCAAAGATGC
AACAAGGTGTTTGTTTTATTTTTTCAAAATGGGTTTGACAGTCTTTCTGCCTACTGAT
TGTCACAGCTAA

SEQ.ID.NO. 106 HG1015008N1 PLT00014333A03.contig.b
 ATGGTGGAGGCCAAGCAGAAAGAGTGACCTCATGGGGACCGCCTGGGTTTGTGTGCCCA
 CTGGAGAGCTCGTTCCTATTGCAGGGCCATAATGTAAAGCAGAGGGTGCTTCTTTTCAGA
 AAGCTCACCTATAGCTCCTTCAGGCCTGTACTCTTAGGAGGGGAAAAAGGAAGGGAGTACA
 GGTTCCTGGAAATTAA

SEQ.ID.NO. 107 HG1015009N1 PLT00014333A06.contig.a
 ATGTCCCTTAAAAAGTGCAGCGAAGTGGGGAAGAAAGGTGTAAATTACTATTATCAGCATCTA
 GAAAGCATCATGAATTGTCTGGAGTACTTCCTAGCACTGACCTCCTTCATTCTGCGTTGT
 TCTTACTGGATCTTTCCATCAGCCCAACAATATGGAAGTACCAATACAAGTCAAATCATT
 CCTGGATTCACTGGAGTTGCTTAAAAAGTTAAATCATTTGGAAATTTTGTGATGATACCTTTT
 CTATATGGATTACAAATTGTATCGCTGGGAAATTCCTCACCTTAAAGAAGACTCTTTTGCTA
 TCTGGAAACCCCATGTCCACCTCTGACTTCCACTCAAAATTGCTTTCCTCACAGTCTGACT
 GCGAGAGTTGTGAAAAATTGGGATGTGCTTCTGAGGTGGGCTGTGGAAATGCCATTACCCCT
 CAGGTGACTACAGATGTGTAAACACCCAGCATGTTCCGGTAG

SEQ.ID.NO. 108 HG1015010N1 PLT00014333A06.contig.b
 ATGCTTAGGAACCACTACAAATGCTACTGCCCTGAGTCTCATTTTGTTCCTCTGGAAA
 CCACATGTGTACCTTGTGCAACAGTATGGGCTCACAGGCAGAAAGGAATTTTCCTTGTC
 TTGGATGAGACTTTTGACTTGGACTTTTGGGTTAAGTTCTGGAGACCAGAAAGGCCAAAAT
 CAAAAGTATGGGCAGGCTTGA

SEQ.ID.NO. 109 HG1015011N1 PLT00014333A08.contig.a
 ATGCTCCAGTTCTTGTACTGGGATGTGGTGGCATCGCAACATTACAGTTACGGCG
 GGATTTAAGAAAGCAGATGAGACCTTTCCAAACCAAGGCCATCTGGCTAGTGTTTGAGCCT
 CAGGTGCAGTTTTCATTGATTTTACCATTAAATTGTGAGATCTGCAATATCAGGTCCT
 CCTTTCACAGAAACAATGTGA

SEQ.ID.NO. 110 HG1015012N1 PLT00014333A15.contig.a
 ATGGGGGCCTCAGCTGCCGTGCAGTTGCCCAAGGCATGGAAGCACAAAGCTCATTTCTGTT
 ACCACATCACTAACTGTAGGATTGGACCTCAACATGCATACTGGTTCTGGCACAAAGGA
 GCTGATTTCTTGTTCCTTGCAACTGTCCAGGCAGGGGTTCTCTCAGCAGATGTGCCT
 CTCCTAGAAAGTAATTCAGGACCAGCCACCTTCCATCCTAGGCCTCCATCACAGCCCCTTG
 GGCCTCGCTGTGTGTGCACCCCCAGCAGTCAGAAAGAGGGAGACTGGGAGCACATCTC
 TGTGGCTGCAGGCGTCACTTTCTGTCTTCTCTGTCTACACTTTCTGTGTTTGTGGA
 CAATCCAGTCGTCAGGCCGTGCCACCTGTCTATGTTGGATGGGAAATGTAG

SEQ.ID.NO. 111 HG1015013N1 PLT00014333A15.contig.b
 ATGGCTGCGCCTCCCTTAGCAAAATCCCCTGCAATTTGGTCTGTGACACCAATGAGGAGCAGG
 TCCCTGTTTGTCTCTGCTGGGCTGGCCTGTCTGGAACAGTCTTCTAGGGTCACAGTGC
 ACGTCTGCCCTCTGGGCAGCTGCCCTGACCCAGGCCCTCCACGAGAGCAGGGGAAGAGGCACAG
 AGACGTTCCATGATGGCTGGGTGA

SEQ.ID.NO. 112 HG1015014N1 PLT00014333A16.contig.a
 ATGAGCAGGGAAATGTTAAAAATAATTCTTGGAGGAAGAGTCTTACCAGCAGGACCGAT
 GTGGCTGGGAAGAGGAAAGCCCTCCAGTCAAGAAGTTGTGTGTGGCTCACCATTTGAG
 TCATCTTTATCACCAAGTGAGATCACCCCTTCTTCTGA

SEQ.ID.NO. 113 HG1015015N1 PLT00014333A16.contig.b
 ATGAAACATTTCCATCATTCGAAAGTTCCCTAGCATATGTTTGCAGTCCATCCTGCCT
 CCTGCTCTTAGCATCATAGAACTACTGATTTGCTTCTGTCACTTCAGGGAAGTTCACAT
 TTTCCAGAAATTTATATAAATAGAAATTGTTTAG

SEQ.ID.NO. 114 HG1015016N1 PLT00014333B03.contig.a
 ATGCCCAGTCAGGGAAGCCGAAGCCTGGAGGGCTTCCGGGAGCAGGGGCTGGAGTTCC
 TCTGCCAGGCAGGAGGCTGGCACCCAGACACCCGGCAGAGGAGCGCGCAGGGCCCCAGCA
 AGGATTCTCCCCAGCCCTGTGCCCTGCGTCTCCTGCGGCTTCTGTGCGGACCGTGTCCT
 TGGGCTGTGTAG

SEQ.ID.NO. 115 HG1015017N1 PLT00014333B03.contig.b
 ATGAGCACGACGCAGTTAGGAGGCTATTGGCCCGCGCGGGGAAAGGGGGAAGGTCGG
 GCTCGGGGTGCAGGACCCAGAGGGCAGGGGTGACTGCCGGGTGCTTTGGGGCAAGG
 TGGGCGTCAAGGCCCCCAGCAAGGTTGGGATAA

SEQ.ID.NO. 116 HG1015018N1 PLT00014333B05.contig.a

ATGCAAGACAAAGTTCTTTTACTCTTTCCTGTTCTCTCCTCAAGCAGAAAGAGGGAGTC
 ACTTTTGTCTGCAAGATGCACTACATGGGTTGGGGTATGGGTGTCAGTACTTCC
 TTAGTTACCTTATCTGGTGTCTCATTTGGTGTCTCCAAGTCCACTTGA

SEQ.ID.NO. 117 HG1015019N1 PLT00014333B05.contig.b

ATGAAGGTGATTTTTTTGTGTGTGTATAGATGGTTATCAAAATTAGCATTCCTGTGG
 GAGATAATGGACAATCAGTGGAGGCCCTCTATTCAAGCCACCTTACTCTGCCCCCTAGATTGT
 TCTTTACAAATACACACATGCTTATATCATTTCTGATGTTAA

SEQ.ID.NO. 118 HG1015020N1 PLT00014333B15.contig.a

ATGGCCATGAGCTCTTGCTGCCATGCCCGCTTAGGAACACTGCTCTTTGGGCTGAGCTC
 ATTAACCTGAAAGACTCAGACAGACAGAAATACITTTAGAACACACTGAAGAACAGTGT
 GAAATGATGTGTCATAGCTGTGAATGGCTGGAAAAATCCTGA

SEQ.ID.NO. 119 HG1015021N1 PLT00014333B17.contig.a

ATGGTGATTGTATTAGAAGGGACTTGGTGGAGGCAAGCCTGCTGCAGCGATCTGTTCCAG
 CCGTGGCGGTGTCAGAAATGTGGCAGTTCTGACGTGGTGGCCCTTTGTGTCTCAGGAGCAA
 AGTGACAAAAATGGCAAGGATGGGGTTGCTCATCTCAACCTCACAAATGAGACGGCACCT
 GGAGAAACACATTTTCATCCCAGACACCTCATTAGCAGAGAGACACAGATAA

SEQ.ID.NO. 120 HG1015022N1 PLT00014333B17.contig.b

ATGTCTCCTACTTGTTAAAAAAGGAAAGAAAAAGATGAGTGCTTCCAAGCCATTGAA

CCACTGCAAGAAGGTTTCTTCTATTTTAAACCCAGATTGTGTATCACAAGGCACGGCTG
AAGGAAGGCATTCTGATAATGCTTTCAGGAATGGTACCAATGAGTAATTTTGTGCTACT
TTCCATTTTGGGGTATGA

SEQ.ID.NO. 121 HG1015023N1 PLT00014333C02.contig.a
ATGGTTCAGGGTAGCTTTTAACTTCATACGACCCCTCACTTCTGTACCTGAAACAATA
AAATGTAATGAATACAGTGTTCAAAAATCAAGTGGCTTGCCTTCTGAGATTCTTGGGTG
TGTCCTCTCTCACATGTAATTTAGATCTTCTCTTTATTATTCTCTCTTACTACCAGCATT
TTTCTTCAAAGTGGAGTGCTCTTGGAAAAAGGCCCTGGTAGACCAATTTTGA

SEQ.ID.NO. 122 HG1015024N1 PLT00014333C02.contig.b
ATGAATACAGTGTTCAAAAATCAAGTGGCTTGCCTTCTGAGATTCTTGGGTGTCCTC
TCTCACATGTAATTTAGATCTTCTCTTTATTATTCTCTCTTACTACCAGCATTTTTCTTC
AAAGTGGAGTGCTCTCTTGGAAAAAGGCCCTGGTAG

SEQ.ID.NO. 123 HG1015025N1 PLT00014333C10.contig.a
ATGGGAAGGGGTGAGGAGCGCTGTGGGAAAGGCTCCTCTGGCAAGGAAAGGATGATT
TCTCCGGAGTTCACCGGACTCCCCCTTCCCCAGCGCCAAAGCATCCGGGTGAGTGTTTG
CAATGGGGTCTGTCAACCATCAGGAACCGAGGAGAGTTTAGACTCTGCATTTACTACACCA
GACCCACAAAGCACCCACCTGTGCTCTCAGTCTCAGGAGAAACAGCCTCTTTCACCTTTTCA
AATCAATCTTCTGTGCAATGTAATTTACTATAATGTTACAGCACCAATGGTTCCAGCTTAA

SEQ.ID.NO. 124 HG1015026N1 PLT00014333C10.contig.b
 ATGAAGACATCGGTTTCAGTTGCCCTGTTTCCGCGAGCTATTCCCTTGTCACTCAACGCTGC
 TTCAGCTGTTTCATGGGACAGCATAAATAACACCTTCCACACTCTGTCTCAGGATGTTTA
 ACTGCAGTCAGGGAAAGGCTGCATTTTCAGAAAGACTCTGGCAGGTCTGAACACCCATTG
 GCCACACTGGTGCA TGGGTGCCCTGAAACAGGTGCAACCCGACTCTTCACTCCCCAGA
 GTCAGATTAGAAATTGGAGGCCAGTCTGCTGTTCTATTAG

SEQ.ID.NO. 125 HG1015027N1 PLT00014333C16.contig.a
 ATGCAAGTTAGCTCTTGCCCTTCTGTCTCTTGCGAGGATTCCCAGGTCAGAAACACAGGAAA
 TACAAGGTGCTGTACAACCTCACAGCTGCTGTTCCGCGAGCCGGCTCTATGGGGACCTGGAG
 GCCATCCTGTACCACGTGCACCTCTTCCAGCCCCACGGAGCTGCTGCTGCAGCAGGCGGTG
 TTCTTCCTGCGAGACACTGAGCGGAGCGGGTCTTCCAGGCTCTAGCCAGGATCCACGAC
 ATCTGCTATAACAGCACCCCTCTGGGACGTGACGGTGAGGGACCTGCTGCCCTCCTCT
 GCTATGATAAAAGACTTGAGCCAAGAGTTTGGGATGCCCTTTTCGCAAGAACTCACA
 GATGAGAAACTGTTTGCCCTACCACTCAGCCTGCCCTCCCAATCTTGAGGACTACCCACAGT
 CGGAACCTCCACCTCACTTAGAGATCCACGCCACCAGGAGAAAGTACCTGCAGTGCGGG
 AGTGCCATGCTTATGAAGAACAAGACAAAGCACAGTTTCATCCAGAAAATATCACA
 GAAGCCTACCAGGTCAGCAAGAGCCTCCGAAGTCCGTGGCGAAGGTGATTAAATTTC
 GCCCCTGCCAAACAAGGCCGTCTACAATACTAGTACCCAGACCAATGAATTCTACAGAGCTT
 GCCAAGAAGGAGCTGTATCAAGAGATAGCCAAGGAGCCAAAGAAAGAGATTACGTACTCA
 CAGGATTACCTCTCAGCCATGGTGGAGCCCCCTGGACTTGAAGGAAGAGGAGAAAGGCC

CAGAGAAATCCCGCCAGGCCTGGCTCACAGCCAGGGATTCCAAAGTGACAGGTCTTCAG
 AGCGACACCGAAAGCAGCTTTCAGGATCTCAAGCTGCCACCCATCAAAGAGCTGAATGAG
 GAGTGGAGGAAAACTCCCTGTTTGCTAATGTA CTGGAGCCTGTGTTGGATCGAGACAGG
 TGGAGCTGGGACAGGCACCACTGGACTTTGATCTGTACAAGAAACCAACCACCTTTCCCTC
 GAGCTGCTCCCTTCGCCCGCACCAAAGCCTGTAAACAGTCAGGAAGAAAGGGAACAGC
 CCCATCTCCTGA

SEQ.ID.NO. 126 HG1015028N1 PLT00014333C16.contig.b
 ATGGCAGCAGCCAGGCCCTGTGTGAGGCTGGGCTCACCCTCGTGGTCGTGGTTGCG
 GAGCCCAATGCGGATGAGCGGCTGGCCCGCGACAGCACGCGCGTCATGCCATACAGGTT
 GATGAGGATGTTGGCCACCCGCTTCAGTACCAGCTGCTCCTCCCATGATGCTCTGGGATCA
 CAGAGGCTCCAAGTGGGACTCACTACCTAGACCAGTCCCCCACATGGTCCCTCCCTGGG
 CTGCATCTTTGCCTGTCTTAG

SEQ.ID.NO. 127 HG1015029N1 PLT00014333C21.contig.a
 ATGTTGAGCTACTGCACACCTGTTATCAACTCCATGTTTAGTGGCGTCATATTGAGAGGT
 GACAGCGGGCCAGCTGGAGTTCGGGTGGGCATGGGCTTGGCGGGCCTGCACCTCGGAGCA
 GCCGGCCGGCCCTGCCAGCCCCGGGCAGTGAGGGA CTGGCACCCGGGCCAGCGGGCTACA
 GAGGGTGTA CTGGGTCCCCTAGCAGTGCCGGGCCACCGGCGCTGTGCTCGATTCTCGCC
 GGGCCTTAG

SEQ.ID.NO. 128 HG1015030N1 PLT00014333C21.contig.b

ATGCCTGAGCCTCCCACTCCATGGGCTCCTGTGTGCTGCCGAGCATCCCCGACGAGC
 ACCACCCCCTGTCTCCACGGGGCCCAAGTCCCATCGACCAACCAAGGGCTGAGGAGTACGAG
 CGCACGGCGCGGACTGGCAGGCAAGCTCCACCTGCAGCCCCGGTGCGGGATCCCACTAGGT
 GAAGCCAAGTTGGGCTCCTGAGTCTGTGTGGGACATAGAGAGTCTTTATGTCTAG

SEQ.ID.NO. 129 HG1015031N1 PLT00014333C24.contig.a

ATGTATATATGTACACACAAATATGCATATGTGTGCACACATATGTGCGCATATGCGGTAT
 GTGTGCACACATATGTGCGCATATGCGTATGTGCGCACACAGTGCACATATGCGTATGTGC
 ACACATGTGCACATGTGTATGCGTGCCATATGTGTGCATATGTAAGTATGCATGTACACAT
 ATGTACATATATGTACGTATGCATGTACATATGTTTGTGTGTATACACAAAAACCCA
 TACATACATTCAAAAATGAAAAGACCCCATGTCTGATTCCTCTTAG

SEQ.ID.NO. 130 HG1015032N1 PLT00014333C24.contig.b

ATGCATATATGTATATGTACACACAAATATGCATATGTGTGCACACATATGTGCGCAT
 ATGCGTATGTGTGCACACATATGTGCGCATATGCGTATGTGCGCACACAGTGCACATATGC
 GTATGTGCACACATGTGCACATGTGTATGCGTGCCATATGTGTGCATATGTAAGTATGCAT
 GTACACATATGTACATATATGTACGTATGCAATGTACATATGTTTGTGTGTATACACAA
 AAAACCCATACATTCAAAAATGA

SEQ.ID.NO. 131 HG1015033N1 PLT00014333D07.contig.a

ATGCGCCTGGGTGAGGCTCCCGAGAGAAGCTGGCTGCGGGCGGGCGGACGAGCTGCG
 CGGACGGGGCGGGCGAAGGAGGCCCGCAGCGGAGGAGGAAGCGGGCGGGCGGAGGCGG

CGGCGCCGGGAAGAACTAGAGGTATTCCTCCGGCGGCTGGAGACTGAGTCGAGCCCGGG
ACCCGAGTCCTCCGGTATCCACGACGCCACCGGAGGCAGTGA

SEQ.ID.NO. 132 HG1015034N1 PLT00014333D07.contig.b
ATGACCCGGGACGGCGCCCCCATCCGACTCTTCGAGGGGGCGGACAGGGCCCCCATCTGTG
TCTTTCGCTCTCGAGCCCCCAGCTAGAGTTGGCTTCAGCGGAATACCTACTGTGCGGGAT
TATTCAACAAGCCGATTGATCACATTCTTCAGCTCTAGCAGTGCAAAGGCTTCACTGAAA
AAGAACAATTACAAATTTTTTTGA

SEQ.ID.NO. 133 HG1015035N1 PLT00014333D15.contig.a
ATGGAGATGGATATAAACTGGTAGGCTTTACTCGAGAAAGGTTCTGGCTGGGCCTTTTTG
TTGGAGAGCCATCAACTTGCAGTATCTTTAGGTATTTTGGAGGGTTGGTAACTTTTTCC
AGAGAGGACTCCACTAATCTTTTGTCAAAAGAAATCTAAGCCTGCTTGCAGGTTCTGCCAG
CTGACTGGGGACTAG

SEQ.ID.NO. 134 HG1015036N1 PLT00014333D15.contig.b
ATGATTGTTATTTGGAAGTTAGACCCCTCCTTGGTCAATCATCTCATTTTATTATTATTTT
TTCTTATTTTCTTGTCTAGTCTCTGGCAGATTTTCTCAACTTTATTTTCAACCCTTCT
TTTGAATTCTTCACTTCCTATATTGTATTTTCAATTCCAAAGAGCCTTTTGTGTTGTTT
GTTCTCTGA

SEQ.ID.NO. 135 HG1015037N1 PLT00014333E01.contig.a

ATGATGCCATTGTTTTTGATAAAAAGAATTAAACAACATAAATGTATGCTAGAAAAAGAA
CTAGAAGGAAATAAGCTGATATATTGGTGGTCTGGTTCCAGAGGAGAGATTTTAGTGATT
CTTATTTCCCTTTATACTTTGGTATTTTCTAAATAATTTCTCTAATAAACCTGTGCTA
ATCTTATTTATTTCAAAATGTTTCTACTTAGAGAAATTTTAA

SEQ.ID.NO. 136 HG1015038N1 PLT00014333E01.contig.b

ATGGTAGGTTATAAGGATTTTCATGTGCTTATTTATTTACCTTTTGTGTTTCATCTGTAT
TTGCTAATTTTCTATTTTATTTTATTTTGTGACAGGATCTCACTTTGTCACTC
AGGCTGGAGTACAGTGGCACAAATCATGGCTCACACAGCCTTGACCTCCTGGGCTCAGGT
GATCCTCCACCTCAGCCTCCTGA

SEQ.ID.NO. 137 HG1015039N1 PLT00014333E04.contig.a

ATGGACACGTATGATATGGAGCAATTTTGAATGTCTTTCACATAGGGGAAATTGCATC
TTATATACCCATAAATTGGGTTTTAGTGAAAGAGCCAGGATGGAACATAATTTCCCTAACT
TCCATCCTAGTTTCCTTTCTACTTGTTCAGAAATGGCCAGTGA

SEQ.ID.NO. 138 HG1015040N1 PLT00014333E05.contig.a

ATGATAAAGGGGTCAATTCAGCAAGAGGTTATAACAAATTGTAAATATATATGCACCCAAC
ACTAGAGCACCCAGATATTATAAAGCAGCTAATAACAAGCAAAATTAATAATAAAGATT
GGAGCAGAAATAAATGAAATGGAAATGAAAAACAATAACAAGATCAATGAAACAAAAA
GTGGGGTTTTTGAAAAGATAA

SEQ.ID.NO. 139 HG1015041N1 PLT00014333E05.contig.b
 ATGGCAGGAAGGCAGAAAAGAAAGGAAGAGAGAACCAAAAAGCAATCAGAAAACATAAAC
 AAAATAGCAGGAGTAAGCCCTTACTTATCAATAATAACATTGAATGTAACTAAAC
 TCTCCAATCAAAAGACATAGAAATGGCTGAATGGATAAAACACAAGACCCCAATGA

SEQ.ID.NO. 140 HG1015042N1 PLT00014333E14.contig.a
 ATGAAGGCTATGGTAGGGACAAATCCAGGATCCTACCAGAGCTTATTACATGGGTGGTCA
 GAAAAGAACGGCTGCCCATGGCAATGGGGCAGAGAAAATTGTACAACAAGCCAAAGAA
 TGCCTATCCTTTTCATGAAGCATTTTGGTGGGGCCATTCTGTGACCTGTCTATGCTTAATG
 ATTCTAGGAATGGATTCAATTACAGGCTGGAATCAGGGCTTCTTTGCAGTAAAGTGACAGG
 CTACCACATACATACCACATATGCACACTGTGGTCTGAACAAGAACACTCGCTTAAAGG
 AAAGTCAGGAGGAAAAGGAAAGGGTGA

SEQ.ID.NO. 141 HG1015043N1 PLT00014333E14.contig.b
 ATGGTTTGTAAGTGGCAGAACTGGGGCGAGAAAGCCAAAGTGGTTGCCCTCCTGTCCCTTCCT
 TCCATTGTACCACAGCGCTGGCCCTCTCCTCTCACCTGACACTCCCATGGGTTCAAGTC
 ATTGCTACCGGATGCCAAACTCTGCAGGTGCTGTGCTGGACACCGGGAAGCCAAAGTGACT
 GGATGA

SEQ.ID.NO. 142 HG1015086N1 PLT00014333E15.contig.a
 ATGACAAAATGTATAATTCTATATCATTAATGTTTATATGACTGTCTTATACAGGCTTTT
 CTTCCAATGGGTTGGATAGAAAATAATCTTTTTTTTTTTTTTTTGTGAGTTGGAGTCTTCC

TCTGTCGCTCAGGCTGGAGTGCAGTGGTGCAATCTCTCGGCTCACTGCAAGTTCCGCCTC
CCAGGTTACGCCATTCTCTGCTCAGCCTCCCGAGTAGCTGGTACTACAGGCACCCGC
CACCACGCCCGGCTAACTTTGTTTTTGTATTTTAAGTAG

SEQ.ID.NO. 143 HG1015087N1 PLT00014333E15.contig.b

ATGTATTACCTTCATGACAGTTTTTCCTTCCCTAGGGAATAATGCCAGGAATCATTAACCT
GAGTGGTTCCTTAATGGAAAACCTGGATTGGAGGATACTTATATTCCTCCCAACCCCCCAT
TCAGTGCCTTATATTGGAGGAGTCAGTGGAACCTTTGTAGGGGATCTTTGTGTGCTTT
GGAGTTTATCTGTTGCTGTTACTATCTGGAGTCAGTGATTGTGGAAAGAACATGTGA

SEQ.ID.NO. 144 HG1015044N1 PLT00014333E24.contig.b

ATGGCTACCAAGAGGTGGAGTGGTCCAGTCAAAGCAAAAGCAGACCAGTCAAAAGCAAGG
GTCAATGGCAACAGTGTTTTTGGGATGCCCAAGGCACCTTTGTTGCTGACTTGCTGGAAGGC
CAAAAAACGATCACATCTGCTTATTATGAGAGTGTCTTGAGAAAAGTTAGCCAGAGCTTTA
GTAGAAAAAATGCCCTGGGAAAGCTCCACCAAGAGTCCTTCTCTACCATGACAAATGCTCCT
GCTCATTCTCATCAAAACAAGAGGAAATTTGTGA

SEQ.ID.NO. 145 HG1015045N1 PLT00014333F07.contig.a

ATGCCAATAGGGTTTGAAAGTTTAGAGAAAAATAGACAAATTCCTAGCAAAATATAAATTA
TCAAAACCTACTCAAGAAAAAATAGAAATCTTTATATAGTCACAAGGTATTAAATATTCAA
TTCTTCCTAGAGACCCCAAAATATATGCTTTTCTACCCGTAA

SEQ.ID.NO. 146	HG1015046N1	PLT00014333G01.contig.a
<p>ATGTCTCCATATGAA TGTCAGACTGGCATCTCAA ACTTAATGTGTCCCCCAAAATCTTG ATTTTCACCCCCAAACCTGCACCTCTCTCAGTCTTGTCATTTTAGTAAATGTCACCTGGT ACTCAATTCTGTTGTTCAAGCTAAATTCTACGAGCTGTTTTTTGATTTATCTCATCCCTTT ACCTTTAAATCCTTCAGCAAAGCCTGTTGA</p>		
SEQ.ID.NO. 147	HG1015047N1	PLT00014333G02.contig.a
<p>ATGCATCCATGGCCGGATGTGTGGCTCACACCTGTATTCCTAGTGCTTTGGGAAGCCAG AGGCAGGCGGATCACTTGAGCCCAAGGATTTTCGAGACCCAGCTTGGGCAACATGGCAAAC CCCATCTCTACAAAAAAATTAGCTGCATGTGTGGTGCGCACTTTTAGTTCTAGCTACC TGGGAAGCTGAGGTGGTGATCACTTGAGCCTGGGAGTTTGAACTTAAATGA</p>		
SEQ.ID.NO. 148	HG1015048N1	PLT00014333G02.contig.b
<p>ATGAAGACCGCAGAGGTGAAAAAAAAGTGTGCGCACACGGGGCGGAGACAGGGTGT GAATACTTGGAAGAA TTCTGCAGTAGTTTTTGAAATTATTGATTGGTTTGCCCTTTTAAA GGTCACCTTATCCTTGTTTTGTGTTTCTAGGCTGTCTGACTCTTCTTACTAG</p>		
SEQ.ID.NO. 149	HG1015088N1	PLT00014333G09.contig.a
<p>ATCGGCTGCGGTGGATCCAGGGAGCTGGGGCCAGATCCACCAAGATCAGGGAAGGTTTC CTGAGGGCCCGCTGTCCCTCAGGGCCTGAAGAAAATCCA TTCTCCGCGTGTGCGTTGG AGAACACAGTGGCCCCCAACCGCCCGCCTGAGACCTCCA CAGGGTCTGTATCCAGAACTT GACGACCGGACGCCATCTCATCTCTCCCAAGAACCCCTGT CAGCCTGCCGGGGACGC</p>		

GGTGCCCGCCCTCCGGCATCCGGGAAACACAGCGTCCCAGCCGGCGCCTGCGCAGGTGCC
ACACTTCCTGGTCTCTCTGGCTGCGTCTGAGGGCCTTAGCTCTGGCCAGGGGCGCTGCA
ACGGGCTCAGCCTTCTAG

SEQ.ID.NO. 150 HG1015089N1 PLT00014333G09.contig.b

ATGAGCTTTACAACACTTCCATCAGAGAACGCTGGAGTGCAAGTGGTGGATCTCGGCTCA
CTGCAACCTCGGACTCCCTGGTTCAAGTGTCTCTCCTCAGCCTCCCGAGTAGCTGGG
ATTACAGGCACGTGCCACCACACCGAGCTAATTTATTGTATTGTAGTAGGACACAGAGTT
TCACCATGTTAG

SEQ.ID.NO. 151 HG1015049N1 PLT00014333H11.contig.a

ATGAAATCCCATAAAGTCAATTTTCTCGGGATTGTGAAGTTCACTGCACACAACCGCC
TTATTTTACATCTCTTCTTTTATTCAACCATCTGGGGACTGGGCAGAAAGTCAATGGTCA
GAGGGGCACATATGTATTGCTTTGTATGATGGGGGCAGAGAAACACAGGAAGAACTGGGA
AATGAAAATTGCTTTACTTCTTATATTCTGACTAGAAAGGCAACCCAAATGGGAATCATA
ACAGGGATAGTTCAGTGGCAAAAAGACACCCCTTAAAAATAGAAAGCTAG

SEQ.ID.NO. 152 HG1015050N1 PLT00014333H15.contig.a

ATGCCTTGCAAAGAACCATGTGACCACCAACACATCAGAGGACAGGAGAGTTTCATCACA
GGATCCTTGATGCTACTCTTTATAGCCAAAGCTGACTTCCTCCTTCCCTCCCTGAGGCT
GGCAACCACAGTATGGTTCCTATTTTCTATAATTTGTGCGCTTAAAGAAATGCTCTATAAA
TGGAATCATACACTATGCAACCTTTCGAGACTGGCTTTTTCGCTAAACGCAATTCCCTT

GAGACCATCCGGTTGGCAAGCATGTCCATAG

SEQ.ID.NO. 153 HG1015051N1 PLT00014333H15.contig.b
 ATGAGCACCTCCTTAGCGGTTGGGGCCGGGCTGTGTCACGGCTGAGAGCCTGGAACCC
 AGCCGCCCTTCAGTGAGGAAAGCCAGGGGAAGGTACCGCGGCAGAGGGGGTGTGTGTG
 TGTGTGTGTGTGTGTGTGTGTGCGCGTGTGCGCGGTGTGTGTGTGTGTGTGTGTGT

TGA

SEQ.ID.NO. 154 HG1015052N1 PLT00014333I18.contig.a
 ATGCTCTGGGATCTGACGGTCCTGGACACCTGTATGGCTGTGTGTGTGTGTGTGTGCCTA
 GCCTGCGGTTACATTTTGTCTGTGGCCACCACCTCCCTCATGTACAAACCGCGTCTCTGC
 TCTGCCAGTCTTGGCCCCCGTCAGGCAGCGGTTCACTCCCTCACCCAGGGAAGTCTAG

SEQ.ID.NO. 155 HG1015053N1 PLT00014333I18.contig.b
 ATGGTAAGTGTGTAGAGAGGGGCTGGAATTGCCCTGTTGGTGTATCTGTGGGAAGCAGCT
 GCTATCCTTGAGGAAGAAAGATGGAGAAGGTAAGTGGCTATCAACCTGTCTCAGAGACAAC
 ACATTCTGTGAAGTTGTACTCTTTCATTTCTGA

SEQ.ID.NO. 156 HG1015054N1 PLT00014333I22.contig.a
 ATGGAGTATCGCCATGTTGCCCAGGTTGGTCTCAAACTCAGGCTGTGAGCTCAAGTGAT
 CCGCCTCCTTGGCCTCCCAAATACTGGGATTACAGGCATGTGCCACCATGTCTGGGCCAC
 AAGTTCATATCTGGAGTAGAAGTTTTACTTTGTAAATATTATAAAGTAGAAGAAACCATA

AACCAATTTTGCTAAAATGAAAGTTGGGGTTAA

SEQ.ID.NO. 157 HG1015055N1 PLT00014333J22.contig.b
 ATGAGATTTTCAGCAATTCCCTTTTGTGCATTTTATTTTATTTATGAGTCTTCTCCTTATC
 AGCGGACAGAGACCAGTTAAATTTGACCATGAGAAAGAAACTGCGCAACACAAATTGCGCTT
 CAGAGGAGATGTATGCCTCTCCATTACGAGTACCCCTTCCCTGA

SEQ.ID.NO. 158 HG1015056N1 PLT00014333J01.contig.a
 ATGAGTGGCCGTGTCCCCCTGGCAGAGAAAGGCCTTGTCTGAAGGCTACGCCCGCCTCCGG
 TACAGGACACCTCCTTGCTCATCTGGCAGCAGCAACAGCAGAAAGTTGGAGTCGGTGCCA
 CCTGGACGTACCTGAGCAGGAGCCGAAGCATGTGTTACTCACAGTATGGAATGAGGCC
 ATCTTGGTCCGAGACAAAACAAGCTCGAGGTCTTAGGGACACAGGGCAGTCCAAGTTT
 TGCACAATTATGTAA

SEQ.ID.NO. 159 HG1015057N1 PLT00014333J01.contig.b
 ATGCGTGCGGGGCGTGCCACCGGAGCTCCTTCCAGCTTGCGTCTGGGGCCACCATGTCCC
 TTCCTCAGGGCACTGGATCCATCTCTTCAAGCTCAGTGTCCAGGCCATGTTGCGCAGAAT
 CACCCCCCAGGGTGTGACGCCCTCGTCCCTGTCAATTCCACACACAGCTGCAACCCGGAAGAGC
 TGTGCGCGGGCGCGCTCTGA

SEQ.ID.NO. 160 HG1015058N1 PLT00014333J13.contig.a
 ATGGCTTTATCACTCGCTTCAAGCCCTACACCCCTGCTTCTTCTTCTTAGAAAAATGGAG

TTATCAAGTTTAACTCAGTTTCTCCTTTCCAAAGTGAAATTGGGAAATCACAAATTTACC
CCTATAACTTACTCATTTCTCTCCACTATTAGGATTCTGTGACCCCTCTCTACCCTTTTTA
TGCTCTTCCCTACTCTTGCTGTACAAATATCTCCTCTCCTTTATTGTCAGTCTTTTCCCC
TTCAGTGACTCTCCTTATCCTACAAAAATGCTCAAGACTCCCCACAACACTACTGCCCATCTC
ATGCAACCCACAACAGTCTGA

SEQ.ID.NO. 161 HG1015059N1 PLT00014333J13.contig.b
ATGAGTCTAGGCAGTGGTTTTATTCTAATGAGCAATTTGTGAGACATACTATTAAATCAT
GCAATACCTTTAACAGTTGAACTTTTTAATTGTGGGTTTGTGTGCGAGGGAGATATTT
ATGATTTCTACTAAAGCCCTCAAAGCTGCTCAGACAGCCCTTTTGTCCATTTTGTGTCAG
CTCCCATGGCTTTATCACTCGCTTCAAGCCCTACACCCTGCTTCTTCATTCTTAGAAAAA
TGGAGTTATCAAGTTTAACTCAGTTTCTCCTTTCCAAAGTGA

SEQ.ID.NO. 162 HG1015060N1 PLT00014333J15.contig.a
ATGGAACCCACAGAGTGTGACATTTATTCCCTGTTCCCTGCAAATTTCTTTATCAAAGAAATA
TACTTTGAGATAAATGTTATAATAATAAGTCCTAGCTTTCCTTAATAATTCTAAAAACAAA
GATGACCAACTTACACCAAGTACCAAGTAGAAATTAAGGTTATAACTACCACGGTAAGT
TACTGCCATTAA

SEQ.ID.NO. 163 HG1015061N1 PLT00014333J15.contig.b
ATGCACAGTGATGGTACCTTGCTCTATTTCCTATATCCATATATCACATTTATCCAATTCACCCAGT
CATCTGTCTTACTCTTTTCTTATCACCAACATAACAAATTTTAGCCTTCATAATGGCTGGA

AATTGGCAAAACAGCAATGAGGGACTACCAATATCAAAACCTATCACAACCACTTTGGT
GTATTCTGA

SEQ.ID.NO. 164 HG1015062N1 PLT00014333J17.contig.a
ATGGGTAAAGATTTGAATAGATGTTTTTCCCAAGAAAGATATACAAATGGCCAGTAAGTAC
AAGAAAAGATGTTCAACATCATTCATCATTAGGGAACACAAATCAAAACCAACAATGAGA
TATGTCACACCCACCGGATAACCACAATTAAAAAGACGATAATAACAAATGTTAGTGAG
AATGTGAAGGAATTAGAACTCTCATTCATTGCTGTATGAATGAAAACTGGCATAACTGC
CTTGGAACACAGTTTTCCAGTTCTTAA

SEQ.ID.NO. 165 HG1015063N1 PLT00014333J23.contig.a
ATGAGGTCTTTCTGTGATTCCAGCAGCCTCAGGAGGACAACGTCAGAAACCAAAAGCTCC
TCTGCCCAGGGATGGGACCCGCCGACTTCAAGGGCAGAACAGGTGCCCTGGCTAGTGGGG
GAGTTCCCCAATGTGGATTTAAACTGTAGTGTCTTCAGAGAACATAACCAACATTGTCT
TTAGGAGATGTGTTGTTTGA

SEQ.ID.NO. 166 HG1015064N1 PLT00014333J23.contig.b
ATGTGTCAGAAATATGTGCAAGCTTTCGACAGGATCCCCGTTTCTCCCCATGGAGCTTTTG
GCAAATGAAGGGTGTGTTTCAGCAGCTGGTCTGATGTCCCTTAAGCTCCCCCACACCAGC
ACACCAGGACTGTCTGTCTCTCCACATTCCAGCATATCTGGGTCTCAGCTTTGA

SEQ.ID.NO. 167 HG1015065N1 PLT00014333K04.contig.a

ATGCCTCAGCCTCCCTCAGCCTTGCCGTCACCTGGCTAAAGATACCAAAGAGGAGACGGAG
 GTTTTCTCCCCCTCCCCCAATGAACCTGGGAAGACAAAGCGATATACTACAGTTATG
 GGACCCTGCCCTTAGGCAAGCAGCATTTGGAAGGGAACTCTTGGCCTGCCCAGTGATGCAA
 GATCAACAAGGCAATTGGGTGCATGAGCCCAATTACATTGAACACTTCTAAGGAAATAAGA
 AAAAGCATTAGAGAAAATGGGGCCCACTAGCCCATTCACAAGAGGATTAAATTGAGGCAATA
 GCAGATAACTACCATATGACTCCATGGGACTGTGCTAGCTAAGCTAAACAACTTTAGGA
 GCGGAGTCAATACCCCCTCTGGAGGGCAGAAATATGA

SEQ.ID.NO. 168 HG1015066N1 PLT00014333K04.contig.b

ATGAATAAGGACCCCTACAACACTACTAGCACAAAGTCTTATTCACCCTTAATTTCCTTAAAT
 TTAGATAATAAATTTCAATCAGCCATAGAAAGCACCTTTTGCTAAACCTCCCCAAGATACA
 AAGCCTTCAGTGTTATGGAAAGATGTAAATAGTAATCTATGGTGCACTCCACATGATTG
 CTAACATGGGGAAGAGGATATGCTTGTGTTTCACATCCCCCTCAGGTCCTCTTGGGATTCCA
 GTGCAATGCATCAAAACCATACCATGGCATGGCTGGGACCCCAATGCAGTACTGGAAAATGAA
 GAATGTGAGCCTGTAGGACCCCGCAGCCCCGGACAATGCAGCTTCCTCGGACAAACACAGGT
 CCCGGATGGGGAAATGTGA

SEQ.ID.NO. 169 HG1015067N1 PLT00014333K08.contig.a

ATGTCCACCATTGTGGATTTCACAGCTACTTGTCAATGGTGAATATTGATCATCATCAT
 TATCTACTGAGCTGCTACCATATCCAGCTACTCCTTGCATGTTGTTTCATTATTTTCTCA
 ACACTCAGCATATTTGCAATATGTTATGTAATATCACAGACAAGGAAACTGAACGCAGAA
 ATGTTTTATTTCCTTGCCAAACATCATCATGA

SEQ.ID.NO. 170 HG1015068N1 PLT00014333K08.contig.b
 ATGTTGGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCCCACCTCAGCC
 TCCCAAAGTGCTGGGATTACAGGCAATGAGCCACCAAGCCCGACCAAAAAAATTTTTT
 TTTTCAATTAGCCAGACATTGTGGCATGTGCTGTAGTCGCAGCTACCATGGGGGCTGAG
 GTGGGAGGATAG

SEQ.ID.NO. 171 HG1015069N1 PLT00014333L13.contig.b
 ATGCACTGCCAAGGCCACCAAAAGGATGACTCCCCAATAACCAAGGGAATCAAGTGGCT
 GATAAGCTGCAAAAGCAGGCCACTCAAGAAACATATTACTTGGGACCTTAATACTCCAT
 TTAAACTTGTTCAGAAATTTAACTCCACTACACTGAATGA

SEQ.ID.NO. 172 HG1015070N1 PLT00014333M01.contig.a
 ATGCCTCCAGGTTCAAGCCATTCTCCTGCCCTCAGCCTCCTGAGTGGCTGGGATTACAGG
 CGCGCGCTACCTCGCCCAAGCTAATTTTGTATTTTATAGTAGACAGGGTTTCACCATGT
 TGGTCAGGCTGTCTCGAACTCCTGACCTCAGGCAATCCGCCCGCCTCGGCCCTCCCAAAA
 TGCTGGGATTACAGGTGTGAGTCATCACGCCACGCCCTCTATTTTCAGGAACCTTAAAGGC
 AGAAAAAACACTTCTCTGACCTTACCATTCCTTAGTGTACTTGCAACCTGGTGTCTCTGC
 ATCCAGATATTCTTACTAATCCTGATGCAATGA

SEQ.ID.NO. 173 HG1015071N1 PLT00014333M01.contig.b
 ATGCGGCACCTCAAAGCTACAACATTAATGCAAGTTACTGATAGAGAATGCTATCAGAGT

GAAAAGGAAAAAAGTTGAAGATTATTTTGGCTTTTTAAAAATAGCAAAGATTACATTATT
TTATTATATTTTGGCTTTTGAGAGTAAAGTTGGAATAATTACTATATACAATTGTTTCTAC
CTAATTATCAAGGAATTTAATCATTA

SEQ.ID.NO. 174 HG1015072N1 PLT00014333M02.contig.a
ATGTTTCAAGAGGATGTAGGTAAGTATGGTGTTCACCTAGCTTTCAGAAAGTCTCATATT
TTGTGTAAAATTGGTATGCTTATTCATGGTTACTTTGGGTTAGTAGAAGCTATACTTCTT
TTCCAAAGGAAAAGACAGACTGTCTTAAGATTCTAACTGTGCACTCTTGGTGATAAGAAAGTTC
TTAGGGTTGTTAAATAGAGCCAGGAAGCGTGAAGCTTTCTCTCTAGATCATGACAACTCA
TCTAGTACCTTGCAATTACAAAGGAGCTCTAAGTCCATTACTTTGCAAGGAATGAAAGTTC
CAGTGA

SEQ.ID.NO. 175 HG1015073N1 PLT00014333M02.contig.b
ATGGAGACAGGGACCGGCTGCTGGGGTTTGGGAGGGGCAGAGATTGAGGAGCCAGTTGCT
GTGCCCTGTTCTGGCAAAAGAAACAAGATGCGTTGCAAGCCAGAATGATGAATTCACCTT
TCACGTCGCCCTCCGTGGCATTTCTGCCTAAATGA

SEQ.ID.NO. 176 HG1015074N1 PLT00014333M07.contig.a
ATGACTTGATGATTTAAGAAATGTAGGGAGGAGTTTCAATCTTAATCAAAATCTTCCTG
ACACAATGGTTCACCCCTGATTTACCAGTTTTTGACTACATTGCTGAGCAGAGAGAAAATT
GCAGAGTTGACTTGACGATATTTTGACTGATTCATGTCCTAGTGCTTCGGGGAAAATTGCTT
AAGGAAGTTGGCGTTGCTCAAAAGTATAGTTGA

SEQ.ID.NO. 177 HG1015075N1 PLT00014333M07.contig.b
 ATGACCCCTGCTTGGCTTGTGTCAAATCAATCTTGGCCAAAGGTGGCTGGTTACCCAATAT
 ACACTGCCCATGAGTGAAAGGCAATTATCACAGGAGGGAGATGGATCTTGGGTGTGGGA
 AACAGGTGTAGAATTCTACTTTTACAGTTTGGTAGAGGTAATGTTTAGAACCTAA

SEQ.ID.NO. 178 HG1015076N1 PLT00014333M15.contig.a
 ATGGCCAAATAAGCTCATGAAACATGTTTAACATCACTAATCATTAGGAAATGCAAATC
 AAAGCTAAAATGAGATACCACGTCACTCATTAGGTGTTATTTTAAATAACACAA
 CAACAAAAATAAGAAATGTGGCGAGGATGTGAAGAAATTGAAACCCCTTGTGGACTGTT
 GTTGGGGATGTAAAATGGCGCAGCTGCTGGGGAACACGATATGGGAGTTCCTCACAAAAAC
 TAA

SEQ.ID.NO. 179 HG1015077N1 PLT00014333M15.contig.b
 ATGTTACAGCTGAGCCAAAGGTGGGAGTAACCCAAATGTCCATCAGTGGATGAATACT
 GGCAATAAAGCTGTGTCTGGAGGCCACTCCGTTTACACAGCTCTCCACACGGGTGAGG
 GCACAGCCTTCTCGTCTCTTTCCTATGCTGCATCTTCAGAGCTAG

SEQ.ID.NO. 180 HG1015078N1 PLT00014333N05.contig.a
 ATGCTCGGTCTCTCTTCAGCCTCAGGCCACCTGTCTTCTCGAGAGACATTGCAATTTT
 GCGTGTCACAAAAATATCTCAAACATGTCCCCAACGAACTCATCCATTCTGAAGCTGCC
 TGGCTTAACGAATGGTGCCATTGTCTCCCCAGAAAGCCTAGAAAGAAATGTGGGAGCCTCA

GTCCATGACCATGCCCAATGCTCCCCACCGACCCCTCGTCTGA

SEQ.ID.NO. 181 HG1015079N1 PLT00014333N05.contig.b

ATGACTATGACCATGCCCAAGGCTCCCTGCTGCTCCCTGGCCCCAAGCTCTGGATGTCTCTG
TCTGAGACTGCTTACTCCCAGCTGGTCTCCCTGTGTTCACCATGTCCACCTTTAACCCCAT
TGTCACACCGAAGATCCCAACCGTAAACCTGATCCAATCCTGTTTAAACTCTTGCGTGTT
TCCCATTTGCTCCTTGCTGGAATCCCCGAGTCCTAA

SEQ.ID.NO. 182 HG1015080N1 PLT00014333N11.contig.a

ATGGCATGTAAATCTAGATTCACTTCTAGCTGGAAGGAGCGAGCCAGAAATTCAGAGTT
GTAGTCTCTTGGCCTGGGTGGATGTGCTACAACATGAACCATAAACCTCCATCACTTATC
TTGGCGGATAGAAAATTGGAACTTTGGAGGAAAGTTTCGCATCAAGCACTCAGACCCCAAT
AAATACTCCCACTCCTTGCAGAGCTACCGAACAATTCTGCACCTCCATTTATCAGGCTACCC
AAGGATGCTTGGAATGTGCTGGCCGGAGCCACCATCAGCACTTGCTGA

SEQ.ID.NO. 183 HG1015081N1 PLT00014333N11.contig.b

ATGATCACTGCTCACTGCAGCCTCACCCCTCCAGGACTCAAGCAATCCTCCCCACTCAGCC
TCTCAGGAGCTAAGGACTGGCTCAATTGACTATAAAGAAATCGAGAATGTCAGCTGACCCAG
GCAACCGAGAGCGCTTTCCTGACTTCCACTATGCACGTGGGTGCATAATTGTGCTGT
GAAGTAATGAAGAACGTGCTTGCTCTGTAA

SEQ.ID.NO. 184 HG1015082N1 PLT00014333O03.contig.a
 ATGCTATCCCTCCCCAGGCCCCCACCCGCAACAGGCCCCAGTGTGTGATGTTCCCCGCC
 CTGTGTCCAGGTGTTCTTTGTTCAATTCCACCTATGAGTGAGAAACATGAAGTATTG
 GTTTTCTGTCTTGTGATAGTTTGCTGAGAAATGATGTTTCCAGCTTTATCCATGTCCCT
 GCAAAGGACATGAACCTCATCCTTTTATGGCTGCATAG

SEQ.ID.NO. 185 HG1015083N1 PLT00014333O03.contig.b
 ATGAGCTTTGATTTAGGGAAGAACATTGTTGGGCACCCTGAGTCCACCTGCCACCAGCCG
 CGGAGGGCATGTTATAACACACCAGACACGCCAAAGGATGGGAACATTGGACCTAATTTA
 GCAACAACAATCCTTTCCCTCATGCATACITCTCTAATTACAAATGA

SEQ.ID.NO. 186 HG1015084N1 PLT00014333O10.contig.a
 ATGCTGTGTTTCGTGATAATATATACAGTGGTTGCTGGAGAAGGGAACCTCCCAGAGGG
 ATCGCTGACTTTAAGCTGCAACATAGCTCAACGGTCAGAGGTGATGTTTCCACCCAAAT
 CTAACGTGTGATTACTGCAGGCTACAAACTTATTTGCCCTCATGGAATAG

SEQ.ID.NO. 187 HG1015085N1 PLT00014333O17.contig.a
 ATGTTTGGAACCAAGGTGCAGTGGCTCAGGCCTGTAATCCCAGCACTTTGGGAGGCAGAGGC
 AGCCACCTCTCAACAACCCATGAAGTGTCTGGAGCTCCACCCTCTGCTCAAGAGGAACGG
 CACATGACCCAGGCCTATGCCAATCAGCAATTCACATTCCTCATGTTGGTTCAAGCTT
 TGTATCTGGGTACACCTGTTTCCCTCTGAGACATGA

Appendix B

SEQ.ID.NO. 188 HG1014903PIPLT00014330A02.contig.a
MHVHVHCSTIHNSKIDSTQMPINDALDEENMVYLHHGILCSRKREQDNVLCKDVGGAGS
HYPQQTNTGTENQTPHVLTYKWEPNNENT

SEQ.ID.NO. 189 HG1014904PIPLT00014330A02.contig.b
MMHWMKKIWIYITMEYYAAVKGNKIMFFARTWVELEAIIILSKLTQEQTKEHHMFSLISGS
PTMRTHRRHREGNNTHLGLSWGWAREGEH

SEQ.ID.NO. 190 HG1014905PIPLT00014330A08.contig.a
MRFPISLHPHQNLLSVFFILDVLEGVEWYLIIVVICISLRTNAFEHFFMCLLAICISL
EKCLFKSFVHFLIGLSFYCCMV

SEQ.ID.NO. 191 HG1014906PIPLT00014330A08.contig.b
MFIQILCPFFNWIVFLLHHGISSLYILDTSLLLVIPFANIFSHSVRCLLTFLMCPLKHKS
F

SEQ.ID.NO. 192 HG1014907PIPLT00014330A17.contig.a
MGDSLALLAKKMKGLDRAGKPPDPADWHWWFLLSTVSTIADGWVFQFENILLPSSSLQFC
VNFSKI

SEQ.ID.NO. 193 HG1014908PIPLT00014330A20.contig.a
 MTYLEIFIQRIYYSIFQCATAATTACFWSECSATNIAILLGKRSAGCWIVEVAHI

SEQ.ID.NO. 194 HG1014909PIPLT00014330B02.contig.a
 MQVHFECNKAHKSFPDFVYMTLNLLSDESSVIRFQSSLAQARKWACMTDQAQSGRRPTC
 CVLSCFPPIAQEDKSSVLGEAKLFS

SEQ.ID.NO. 195 HG1014910PIPLT00014330B02.contig.b
 MESHSAQAGVQWCDLGSAAQALPPGFTPFSHLSLLSSRDHRRPPLHPANFVFVFLVETGF
 HCVSQDGLNLLAS

SEQ.ID.NO. 196 HG1014911PIPLT00014330B04.contig.a
 MRIPELGORKQAGRKMGTCGAGGAGGDLQLGDPQHSTCPALCLVQGCNLPPASPRLHR
 APPGEGLVVHTQHLCSLERMGCETPDASQLPSLERLVEWHISLGSLPRVPSAPAVHAVG
 PHPSGKRSLLAWVVFPIRNCQVLGLDGLFPISVGEGGIV

SEQ.ID.NO. 197 HG1014912PIPLT00014330B04.contig.b
 MGPAALARGDPFSRGLSSPWGPARRKSARLPNLGHTLSAPLSPRRNSRSPGRQGP GS KS
 PEFPELGWARGPLMRRLPGASQPQLIAGPDQSRGERPCSLPADRDL

SEQ.ID.NO. 198 HG1014913PIPLT00014330B05.contig.a
 MRSNLITDILHSNRYYTLYSRILRQQPYKVIDAETGPKSKNNLFRVFQPSGEGETWVC

LKPIFTIFFSYDTLPITNE

SEQ.ID.NO. 199 HG1014914PIPLT00014330B11.contig.a
MAPAAAPSSLAVRASSPAATPTSYGVFCKGLSRTLAFELAWQLRMNFPYFYVAGSVIL
NIRLQVHI

SEQ.ID.NO. 200 HG1014915PIPLT00014330B13.contig.a
MSSPFTHSQDPETSTAYKTLHQWPPTFLAPGTSFVEDNFPDWDGGVGAVWGCEM
MEFYRSDILTEVYCKIRYSLREKRRQFRGQVERKYTDKVCRSAQGSEAVSWKT

SEQ.ID.NO. 202 HG1014917PIPLT00014330B18.contig.a
MLFFPSHSILTLSILRSQSCREAAFPWQPIICVIVPCCTIELDTGAMLCLEMYPASK
GLPQSPLLKKPQFY

SEQ.ID.NO. 203 HG1014918PIPLT00014330B18.contig.b
MLDVDRDKWLTGHSSCLSCYGLSSLIQSSLISLTVLSLSPVDSL VGHWFLSVLN

SEQ.ID.NO. 204 HG1014919PIPLT00014330C06.contig.a
MFVVCGMFCYCGPSCWTQGHWPAPRHHPSLTLLSWNTVVPHPCKATSRLLHTPCFFH
SMVKFFSYQEAGQVQWLTPVPALWEAETGRSFEVRSSRPA

SEQ.ID.NO. 205	HG1014920PIPLT00014330C06.contig.b	MTVSRAPHVLDPRWAEATKGDPARVKGGCRQARAEPICRMLRARAPPSTGSPSELSIDWT
NHPQC		
SEQ.ID.NO. 206	HG1014921PIPLT00014330C12.contig.a	MEPFRRAPQWGIYDPWFPSRKLQRKKRAGQPMRERRQRKRKGDGRKGN AFFRTQEKQH
QKWAFLPV		
SEQ.ID.NO. 207	HG1014922PIPLT00014330C14.contig.a	MQLSNKPSQPGLKVSNSATIKAILCHDQEHRAWSQTDKGLRLMKDCQVTFWSLLHSLEIR
TWSFKL		
SEQ.ID.NO. 208	HG1014923PIPLT00014330C18.contig.a	MPGQRGPAPASGWKGREQRGGSYSGPPGQSPGVGKQAKRRPLLSSLTPRPEPKVQVPQ
AQSP		
SEQ.ID.NO. 209	HG1014924PIPLT00014330C18.contig.b	MSTLSSAGPQLEGAIQSSFPSPAELQRWGGCRRAQGCLGRGLCA FCCGGHGHQFRKILNGL
SID		
SEQ.ID.NO. 210	HG1014925PIPLT00014330D03.contig.a	

MTPLRWYLCTCLFKATFSFLSWTVTDPLYPKNTPLDFRAGCLAYTPGSLSQGCACH
CNKEEKTSHGKGLTPSYNQNTQLHREGTPLSHFGSHIHHSVTPTQAGSQVSGVLVLTRAK
RTKQDPRSPFKH

SEQ.ID.NO. 211 HG1014926P1PLT00014330D03.contig.b
MCISLLRFHMYSFIMRIFFLHPMTIVLIAALKYLTADNNILDILGIAFNAYVLSVCVWTF
LCFFSCLLNFKIVF

SEQ.ID.NO. 212 HG1014927P1PLT00014330D05.contig.a
MNTRSGKPLETGYHGHLHTNMKLSAFPPTTYTIFNVFPVMDFKKEVIISNNIHEIMDIM

SEQ.ID.NO. 213 HG1014928P1PLT00014330D05.contig.b
MGFPILVSQFRSWYCLSIISYPLLGNKLLPNLAASNNNKHVLCIVCFRTLGVA

SEQ.ID.NO. 214 HG1014929P1PLT00014330D07.contig.a
MVCVIPDLFSCIKNIRMNLKVYCHRKSTDQNPGRKSGKSKGLALTSNSTAYKDKGG
SILSFSIIGPYTLRFCWGVANSCLS

SEQ.ID.NO. 215 HG1014930P1PLT00014330D10.contig.a
MFLVEMGFHHVGQAGLQLLTSSHTLASASLSAGITGTSHRAWPGICIFASNPGHLSAHQS
LKPLLGGCVLASRLPGPMD

SEQ.ID.NO. 216 HG1014931PIPLT00014330D10.contig.b
MNSLLTLVLFVCLFSPFLVKCANILKSQENHMOVISHSPLWPASAKMTFCHSYSIHFCHAT
PDIHTCSFWSIK

SEQ.ID.NO. 217 HG1014932PIPLT00014330D12.contig.a
MRKMTFLSADGGGNKNHRDCDFLMVLWVMVEFKLTTKFLVTCFLVFTENILFFETESLTV
SPRLECSGVMCRLLTATSASWIIQVILMRQPPSSWDYRPPRPANFCIFSIDGVSPC

SEQ.ID.NO. 218 HG1014933PIPLT00014330D12.contig.b
MEFHVVDQAGLELLASSDPPVVSASAGITGISHHTQPKTFFIMIYQYQTRNY

SEQ.ID.NO. 219 HG1014934PIPLT00014330D13.contig.a
MKATMSLLNKSFIKRNIYQGLPCHIDPLKVRSEVEAALMFLILNLIQQKKSKDILTSIDF

SEQ.ID.NO. 220 HG1014935PIPLT00014330D15.contig.a
MANKHMKKCSNFLVIKECKSKYEIVLYFTRTGMAVIEKTDNDNKYWRGGEFRNSLVVFQK
INIESSNLTPILDIYPRAMRTYVHTKVCTF

SEQ.ID.NO. 221 HG1014936PIPLT00014330D15.contig.b
MYDHGLKKLFFFFFFFFETPCFVAMLGCSGAILAHCNHLPGSRDSPASASRVAGTTGT

RHHAQLFVFLVEMGFHHVGDGLDLLTS

SEQ.ID.NO. 222 HG1014937PIPLT00014330DI7.contig.a
MGCPSIAEMHQGHISHLLCLGCPICIYQRKPWTPTRGASMRECMPLYKFTPTSEKRPQL
MLPLPEQCEQLCRFGSTPVTWALJWFGCPTQISS

SEQ.ID.NO. 223 HG1014938PIPLT00014330E04.contig.a
MLDFTHRSGFRKKQDASAVALYTITLKDRLSCFCFISPYQLPPCCTLQNFGNKT

SEQ.ID.NO. 224 HG1014939PIPLT00014330E14.contig.a
MMLETHQDCTLEAAVTSEERCPLLWLVKRNQVRLPERFGPLQPATAGLAMPSPWRNRD
STATEITY

SEQ.ID.NO. 225 HG1014940PIPLT00014330E14.contig.b
MELKNEGEMGGQGTNWGMTAYSVRIGNAHKGQMKMISLCVIRRIATKIRVKGEGGKGR
E

SEQ.ID.NO. 226 HG1014941PIPLT00014330E24.contig.a
MHFGGHIQPISTTIYLGSDFTMENLFLPQDRQDWKTFPGAAWESPAHQGTGCLSSH
PTQARELTPESQGWAAQQPPCSPLAPAVSHLTGPRGQGSFHSNYLSYIYTLPK

SEQ.ID.NO. 227 HG1014942PIPLT00014330E24.contig.b

MYTYIYVCISDYKCTYMYIRFHIYCYLFICNEMSQKRNNKKKKRKKKKKKKKRKEKK

RK

SEQ.ID.NO. 228 HG1014943PIPLT00014330F01.contig.a

MESWIAPPFA YEHGYIFSTIFHELKIRLVILHIFFLIYFWIFLLFQGPRVWSDMIELTVG

MNQRQ GAYKVSRA TIHC

SEQ.ID.NO. 229 HG1014944PIPLT00014330F03.contig.a

MQQRQTMVWPKMEQN LQGKREVPGVICVGQMRMELVNEQKRRLEECSGDDSPPLLLPQRK

ELQEILNWLKDTQDVL FVCHL FVLLMTPADTTLASSGNKCCCLQHL

SEQ.ID.NO. 230 HG1014945PIPLT00014330F03.contig.b

MEGDSFELCCRSAIAYAVCKVTPSLGASVPFCSPFLDVPCCSLRSFSASSLMWLCSVFC

LMPQALVPATK

SEQ.ID.NO. 231 HG1014946PIPLT00014330F04.contig.a

MNMLYIALLLPPLTMLEESPTGCLHRTHTTWSGNSITKTLLYHTYYGCMGNRLGTCTY

SQTTYSVCDPGNNQLYVCYDPKFPGEWFEIRAVKRRSPLKPNQGPSLLLRDYFSVF

SEQ.ID.NO. 232 HG1014947PIPLT00014330F04.contig.b

MRNAIQNRLALDYLLALEGVVCGKFNL TNCCLEDDNGKA MEITARMRKL AHVPVQTW

KGWSPDSLFGGLVFIFLRVQDFNRS GSGHIRKLPNTPLSLTSPC

SEQ.ID.NO. 233 HG1014948P|PLT00014330F05.contig.a
MSVIQTLWEADAGRSLEVRSSRPARTWQNPVSTKNMKKKLGWCAGTCL

SEQ.ID.NO. 234 HG1014949P|PLT00014330F13.contig.a
MGRSRCPSPRRRLVVSARAVSPVAGKTGRPASSVPTTQPLRPQLEPGSRRG

SEQ.ID.NO. 235 HG1014950P|PLT00014330G21.contig.a
MVPHHLSSDRLLFLTSVSMFPHDTLSFSSMDTSLTITTALEKLGPLLGSHPKPHTNF
LRSSPSAPFSANTPTSILSEHPVLVSTVLLSPFQAVPKVIIQGYVSFHLKCIIPPLLGSQ
ISFPGPHLINFLNRIIINNDTSVF

SEQ.ID.NO. 236 HG1014951P|PLT00014330G21.contig.b
MYICILCIYNDEKTDIFTLTYTYVCVYLWTKPIFSMNCLFLLISYRICMLHNI

SEQ.ID.NO. 237 HG1014952P|PLT00014330H05.contig.b
MNSRFESVTKCLPKEKSPRLGFYADFYPIYKEKQTPILLKLFPKIEEGILPNSLYKASI
TLISNQDKDTTKGENYRPFLMNTDGKILSIILPSQI

SEQ.ID.NO. 238 HG1014953P|PLT00014330H06.contig.a
MNIEMEIGFSPITNFVPQFLKSGMFNSASHLARVCNKQLVIFHPLSTQN

SEQ.ID.NO. 239 HG1014954PIPLT00014330H12.contig.a
MGLLPQAGLAGLGGWGAARPSLKAPRGAVGRAQGRSVVLLVPSCGAGSRAVHGRGRDPA
PRGRSPAPARPWPRPRGGGAQSRP

SEQ.ID.NO. 240 HG1014955PIPLT00014330H12.contig.b
MRRSRTRSLADKPVESAILERVPEAPGFQFRLLTRQPCDLGQLNVRKLGSVFSIGLMSV
RGSRKGGWPAGNLRSM

SEQ.ID.NO. 241 HG1014956PIPLT00014330H14.contig.a
MLVIIRLSCKTALYGLRDYCLSYPLLGTIRLSVSPEVVQJNVVVRHKPREGKVCLNS
FTIMHGNV

SEQ.ID.NO. 242 HG1014957PIPLT00014330H14.contig.b
MSMCVCYNFSFKKSTPGSLSTFLNFSFAHNFRFTEELQNSSYMPFTQIPLVLVFYMTMVY
LSKLRK

SEQ.ID.NO. 243 HG1014958PIPLT00014330H18.contig.a
MALLCHIFLFTVLLPPFTLTSLPPCCCTTSSSSYQEVLWRMWLPRNIDVPSYRGFSKGGP
TFTTHNHIPLHFRPYISIPVSLTSLSMSLPESKL

SEQ.ID.NO. 244 HG1014959PIPLT00014330H18.contig.b
MRIKRNNRHKRPDTGNNIFDKYEFLFFPLFYFTATPTSYSLSFATKQTSGFVCLFHQGEV

LNVSFNPQSIHSLRG

SEQ.ID.NO. 245 HG1014960P1PLT00014330I11.contig.a

MLSKNIRTCETLFVYTQTHLHIYLYLSIEIYCVVCVCMYLCVCVSLSVLVHLSYNT

MP

SEQ.ID.NO. 246 HG1014961P1PLT00014330I12.contig.a

MSLRVLLSSRHILYVSPLEHLPKPLFRLTIRFWSVGPVCILCIPRYSAWDQPELEIQ

LLEHFKTHVEGRDSTYKSPQPMSTL

SEQ.ID.NO. 247 HG1014962P1PLT00014330I12.contig.b

MCLVLQFSTSFLSAAVFLFSFLHKLRLWAQLEKILKPSLALLHSNIFKKCFTGWFF

VVAFLF

SEQ.ID.NO. 248 HG1014963P1PLT00014330I13.contig.a

MAAPAPRRRVPSPPARVPGERWSGVGPVCLPPAPSPNPGLLVKGRGRARARKGGGSEKQ

RGPAERVPTVTPRSSQGAPAPKLLARRAQSRFPDQSRGRRSQ

SEQ.ID.NO. 249 HG1014964P1PLT00014330I13.contig.b

MHSDLMCIFSLFKRIGHLMKGRTRRDDKRRLLACSCFQGGKKNKVECLQFLQNMVAFVISEF

LYLAYKGGGLFRETRMEPFRDMSTL

SEQ.ID.NO. 250 HG1014965PIPLT00014330J10.contig.a
 MAFLQFFLSKPMARAKWEKSETLKGEPIYPKAQRTKAMWVQWICSRSPGLCPRVSENHP
 FWQGKKVSALWMKLSVRPPLLLTPSTAGAGRFQATGLHRPSMLTCTPAHSLTNWSLPPHLL
 ELTRASKADG

SEQ.ID.NO. 251 HG1014966PIPLT00014330J10.contig.b
 MGDPRPSLEYRVSFNLLKTGEQSSQAKLPPAPPQQTSPWTISPAPCRRILQPQCQFY
 RCYHPARKDFEKSSSVLWAVGSHLHSPRPQRKAVSLCRAEV

SEQ.ID.NO. 252 HG1014967PIPLT00014330J14.contig.a
 MLFSIVGEPFYPPQQVHHFILHSGCTILVPANSAQWLQYLLLLTNTCYLLFLDNGHLSRC
 EAYLTPYTKNSKWIKGVA

SEQ.ID.NO. 253 HG1014968PIPLT00014330J14.contig.b
 MFATWHSIANYIHGAVQHASRALPSHVAETLHVPPPTPGNHYSLLCVYKFDCFRDLI

SEQ.ID.NO. 254 HG1014969PIPLT00014330J15.contig.a
 MKEIHRKGPLSEQNLVINFIAMKFPRAFVVFQTDENLCIILNPATCVQYAYFVLKKNVSF
 QSVKNIQK

SEQ.ID.NO. 255 HG1014970PIPLT00014330J21.contig.a
 MGSIRNNLKCIIPLTYWFPSTTLFCRYTHIYIHYSIVYGSKKLETMAGTTKKRDWLNKQK

HSYAVECYAAGRKKGQIDMY

SEQ.ID.NO. 256 HG1014971P1PLT00014330J21.contig.b
MENGWVWWPMPVIPALWEAEVGRSPEARSSRPAPWPTWQNSISTKNTKPSMV EHCSPSYL
GGCGTRIT

SEQ.ID.NO. 257 HG1014972P1PLT00014330K01.contig.a
MALAFFHMSPTNFTELTPTPTPPGMEARRDSSIVQKSGENKKGEGDIAAVWCFIL
CVCFLPRKKNRL

SEQ.ID.NO. 258 HG1014973P1PLT00014330K08.contig.a
MCHHVLLVFKFFCRNRISLHCPGWSPSLEQSSCLSLPEWWDYRCEP THLAFFCLPFNSVC
CFLTYRGFKFCCSWITLPFVTFEIASIKLNLFPFQIQK

SEQ.ID.NO. 259 HG1014974P1PLT00014330K08.contig.b
MKSSIFFHINWPFIFLCELSVITFYLLHSNTIQNIILLQLFYFLLTFLFI

SEQ.ID.NO. 260 HG1014975P1PLT00014330K09.contig.a
MAKRGQGTVLAIASEGASSKPWQLPHGVGVSVMQKTRTEVQETLAGCMKMPGCSDRSLQ
GWSPCGRTSARAEQKGNVGLKSPHRVPTVALTSGAKKRGL

SEQ.ID.NO. 261 HG1014976P1PLT00014330K09.contig.b

MGRYYKDTKNMEATLELGNRQLKQFEGLRRRQENGESLELPRNLEDSEDREMWESLELP

SEQ.ID.NO. 262 HG1014977PIPLT00014330K15.contig.a

MMPGETHSAAPGTAADLRCQCASLQQNLNEYVEALITLKQKIINTDNLLEYQKKCDD

ILLAFSVEFWN

SEQ.ID.NO. 263 HG1014978PIPLT00014330K15.contig.b

MAGRPQETYNHGRRGSKHILHMVAGERSAERSGEKPLIKPSDLVGTHSLSGEQHWGTAPV

I

SEQ.ID.NO. 264 HG1014979PIPLT00014330K24.contig.a

MSVKGAKMLNKYYFQIFFSSLFIPILFSPLVNFIFLKILPFFSKPYISPKKE

SEQ.ID.NO. 265 HG1014980PIPLT00014330L01.contig.a

MPCFKIFFSHTVHLSTRIPKHFSLVFINTPVCVLSQIKNSNFWPGAVAHTCNPSALGGR

GISTKTTKKLAGCGGGLLWSQLPWRLREENGVNPGGACSELRSRHCTPTWA

SEQ.ID.NO. 266 HG1015004PIPLT00014330L24.contig.a

MQGCGRGRWDLHCDLANVDCADCLLRASVSSSGIQSSCLRDCWIRGDGS

SEQ.ID.NO. 267 HG1014981PIPLT00014330M02.contig.a

MVSFYPLESHCLKGHHAPFFQPYLVLQMPRRYL VVWPSLLQGLSPLGHPPTSTSSWLNP
TQTSKMKRKATFRKPSSIQCFFVSSSSPLKGLRMMSSKFDHLGF

SEQ.ID.NO. 268 HG1014982P|PLT00014330M02.contig.b
MERLPHPVLYTTTLQGLSNWAFNFSLIQVSSIGQLLGNWFTHADKPSSPELTHSPSLTTE
EQIRKQDLSGAMTSSSTDGKEMRGAATC

SEQ.ID.NO. 269 HG1014983P|PLT00014330M08.contig.a
MYPCVLIQHPLIGENMQCLVFCVSLLRIMASSSIHVPAKDMISFLFVATQYSMVVYMYHV
FFIQFVIDGHLG

SEQ.ID.NO. 270 HG1014984P|PLT00014330M08.contig.b
MSSTMVELIYSTNGVKVLLFLHSLASIFCFLTFYFIYLFYLFCTPAWATQ

SEQ.ID.NO. 271 HG1014985P|PLT00014330M15.contig.a
MGENKGERSETAACARPPLAHSRPAAPRPPPSLPRLLLTPTSERPGRESGG

SEQ.ID.NO. 272 HG1014986P|PLT00014330M17.contig.a
MAAAA VRPVVTWTTGLGPTTPPRLEDAQRRLGLGRGVSQREPRSPGRRRLGAGSLAGRP
RLPPESAGAPRRPCPWVHRPDPRVPRLGLLLRITLTSFLSRVTHSSSLGPF

SEQ.ID.NO. 273 HG1014987P|PLT00014330M17.contig.b

MLESSKVFNQASGMISFLFCFVFFETESGYIARLECNGAISPHCNLR LPGSSDSPASAPR
VAGTTGARHHAQMRKLRHREVK

SEQ.ID.NO. 274 HG1014988PIPLT00014330N10.contig.a
MKPNPSKEVKSRFITVPSKFLFGFSLWSFISLPFSLCLLSTLSSLSLHHHTHTHTTYV
LSGMSYCQGEGLPRC

SEQ.ID.NO. 275 HG1014989PIPLT00014330N10.contig.b
MLTHTHTHTHTHTHALGGGKLEKTFGARENNEGQVWVFAPVIPAHWEAEVGRLLLELR
SSRPAWAT

SEQ.ID.NO. 276 HG1014990PIPLT00014330N12.contig.a
MSQQSNWEPNSDQQGGIVWLIGSDRNLGRKQPHTKDNKKVTWALKDSVKKPETPK

SEQ.ID.NO. 277 HG1014991PIPLT00014330N12.contig.b
MIKKGKCPLAASTVRLELTEKELLYYLLHSLQRKLEKWVDVIERTCVCGARQTWV

SEQ.ID.NO. 278 HG1014992PIPLT00014330N13.contig.a
MMTMLPVFVLTYCLLFRVYCSYLSKKLTAKQLQAGPSGDVPEEGIAIDDDSSMHVIA
PEELSAVQDVEVEDSDIDDPDLV

SEQ.ID.NO. 279 HG1014993PIPLT00014330N13.contig.b

MGSAlyKSITFFyLVyHMFtVPPCLAMIRyTNSyHRGTTAYsIQYSHMLRRFVA

SEQ.ID.NO. 280 HG1014994PIPLT00014330N22.contig.a

MSDAFSRQILAAELEITKIIMTALTAVPRLMKNAFEQPGLTGIIFANPII GLTSGESIND
FGYVLLCLKDWDLG

SEQ.ID.NO. 281 HG1014995PIPLT00014330N22.contig.b

MGPQSFPLVFR LKATLVAGKSTGKALFQEEARNPGENPSSAPNLQDGLRWISFCTL

SEQ.ID.NO. 282 HG1014996PIPLT00014330O03.contig.a

MWGEVRVIFLGFMAGLGEKGFLFFMTHLGEEGFNYGLLQERMORDERQEGRRRSERETLL
LRSSLWGIIF

SEQ.ID.NO. 283 HG1014997PIPLT00014330O07.contig.a

MPSKTTPISSDDGLCHRPSTEHHDTHRHSSLKTWIDPGATILSSLECW RGPDQGRTPMSG
SRVSPGSWGTLGLAIAT

SEQ.ID.NO. 284 HG1014998PIPLT00014330O07.contig.b

MHKRQQRQVKRRRLVGLQKRPAVPMLRFSHFTDNKSSVSTD TNPGTAGPVPGMTLGQTPWR
WIRKLQFYQGSCC

SEQ.ID.NO. 285 HG1015005PIPLT00014330O18.contig.a

MPEEIEPREWRSREQRRMGSRNRPREPARCFFLSLLENSVQKPLGFRLETERNRLCVF
HCRVSVLGVYSMWVCDQFNCA

SEQ.ID.NO. 286 HG1015006PIPLT00014330O18.contig.b
MEEPGTKEKDGIEQAKRTSSMLLPLPSRKFCSEAIRLQARDREEPPLCFSLSCLSAGSV
LHVVS

SEQ.ID.NO. 287 HG1014999PIPLT00014330P07.contig.a
MLTRLVSNQPRDLPSASQSGIAGVSHCARVPVSFFLKKRVPSSAFILKPPFKTLLL
LSNSEGPFRLDRMSKGQSIAPNTLL

SEQ.ID.NO. 288 HG1015000PIPLT00014330P07.contig.b
MVVCSEKSHLIPPPWHLSDLSCSVGGMFPPLFACSDVFLPLFFRLDCQQAEPGEMKQ
S

SEQ.ID.NO. 289 HG1015001PIPLT00014330P09.contig.a
MNTSAIHDSCFSSWPLKSSLLSLHSYPPQLNFSFPTQEEPVCPLHPTALPDASFPTGM
ASGPHLQVSGPHLQTQHMAFGSSHQESWKGWVEQAEERKM

SEQ.ID.NO. 290 HG1015002PIPLT00014330P09.contig.b
MPPASHCTSGCLFPQDWNGLRPPSPGEWPSPTDSTYGLWLFPLPRVLEGMGRASRGKEDV
SSQNAPHLLSGLLTGIGGLAAQHGPWPWETPSLLS

SEQ.ID.NO. 291 HG1015003PIPLT00014330P15.contig.a
MRLWGECKNLKDKVAQSNQMSIQGRFLSHPLQKSPSRARVNQWKFLALTAHYSYSFINV
G

SEQ.ID.NO. 292 HG1015007PIPLT00014333A03.contig.a
MFRAAAKLGGNIGKIFQEHIGLGMATCIWLSTFLHICGEKKSFFFKIRSHTFILRLLQRC
NKVFCFYFFKMGLTVFLPTDCHS

SEQ.ID.NO. 293 HG1015008PIPLT00014333A03.contig.b
MVEAKQKSDLMGTA PGFVCPLESSFLQGHNVKQRVLLFRKLTYSFRPVLGGKKEGST
GSGI

SEQ.ID.NO. 294 HG1015009PIPLT00014333A06.contig.a
MSLKSATKWGRRCNYYQHLESIMNLLLEYFLALTSFILRCSYWFPSANNMEVPIQGQII
PGFIWSCLKVKSLEFLMIPFLYGLQFDRWEEFSTLKKTLLLSGNPCPPLTSTQNCFPHSLT
ARVVKNWDVLLRWA VECHYPQVTTDVLTPSMFR

SEQ.ID.NO. 295 HG1015010PIPLT00014333A06.contig.b
MLRNHYKCYLESHFVSLWKPHVYLV CNSMGSAEGIFLVLD EFDLDFWVKFWRPEGQN
QKYGQA

SEQ.ID.NO. 296 HG1015011PIPLT00014333A08.contig.a
MLQFSGTDVVASQTFTVTAGFKKQMRPFQPKAIWLVFEPQVQFFIDFYHLIVRSAISGL
PFTETM

SEQ.ID.NO. 297 HG1015012PIPLT00014333A15.contig.a
MGASAAACSCPGHGSTKLISVTTSLTVGLDLNMHTGSGTKGADFLQLSQAGVPAQQMVP
LLEVIQGPATFHPRPSPQLGLAVVCTPSSQKRGRGLGAHLGCGRRHFLLLFSCCLHFLCFVG
QSSRQAVPTCHVGWEM

SEQ.ID.NO. 298 HG1015013PIPLT00014333A15.contig.b
MAAPPLANPLHLVVTPMRSRSLFCLCLGWVPVWNSLLGSQCTSASGQLDPDGLHESREEAQ
RRSMAG

SEQ.ID.NO. 299 HG1015014PIPLT00014333A16.contig.a
MSREMLKNNSWRKSLTSRTDVAGKRKALQSRRLCVWLTIESSLPVRSPPF

SEQ.ID.NO. 300 HG1015015PIPLT00014333A16.contig.b
MKHFHSEKFPSCILQSIPLPALSIIELICFLSLQGSSHFPEFYINRIV

SEQ.ID.NO. 301 HG1015016PIPLT00014333B03.contig.a
MPSQGSRRPGGASGRGWSSSARQEAGTRHPAEGGEGPARILPSPCACVSCGFCARTVS
WAV

SEQ.ID.NO. 302 HG1015017PIPLT00014333B03.contig.b
MSTTQLGGYWPAGEKGEGRARGRTPEGRGDCRGALGARWASRPPARLG

SEQ.ID.NO. 303 HG1015018PIPLT00014333B05.contig.a
MQDKVLFTLSCSLLKQKEGVTFVSARCTTWGWMGASTSLVTLSGVSFGVSKST

SEQ.ID.NO. 304 HG1015019PIPLT00014333B05.contig.b
MKVIFVCVYRWLSNLAFLWEIMDNQWRPLFSHLTLPLDCSLQIHTCLYHSDV

SEQ.ID.NO. 305 HG1015020PIPLT00014333B15.contig.a
MAMSSCCHAPFRNTALWAEIKPERLRQTEILLEHTEEQCEMMCHSCEWLEKS

SEQ.ID.NO. 306 HG1015021PIPLT00014333B17.contig.a
MVIVLEGTWWRQACCSDLFQPWRCQNVAVLTWWPFVSQEQSDKMARMGVAHLNLTNETAP
GETHFIPDTSLAERHR

SEQ.ID.NO. 307 HG1015022PIPLT00014333B17.contig.b
MSPTCLKKGKKKMSASKPFEPLQEGFFYFKPQIVYHKARLKEGILIMLSGMVPMNSNFVAT
FHFGV

SEQ.ID.NO. 308 HG1015023PIPLT00014333C02.contig.a

MVQGSFLTSDPHFCYLKTIKCNEYSVQKSSGLLSEJSWVCPLSHVFRSSLYYSLLLPAL
FLQSGVLSWKKALVDQF

SEQ.ID.NO. 309 HG1015024P|PLT00014333C02.contig.b

MNTVFNQVACFLRFLGCVLSHMYLDLLFIILSYYQHFFFKVECSLGKRPW

SEQ.ID.NO. 310 HG1015025P|PLT00014333C10.contig.a

MGRGEERCCKGSSLARKGMISPEFTGLPLSPAPSRVSVLQWGLSPSGTEESLDSAFTTP
DPQSTTCALSLRRNSLFHFSNQSSVHVFTIMLQHHWFQL

SEQ.ID.NO. 311 HG1015026P|PLT00014333C10.contig.b

MKTSVQLPVSAAIPCHTQRCFSCSWDSIKIHLPHSVSGCLTAVRERLHFQKDSGRSEHPL
ATLVHGVPENRCKPDSSPPRVRLEASLLFY

SEQ.ID.NO. 312 HG1015027P|PLT00014333C16.contig.a

MQVSSWPSVSCRIPRSEHRKYKVLYNSQLFRSRLYGDLEAILYHVHLFQPTTELLLQQAV
FFLRDTERRRVFQALARIHDIQYNSTTLWDVTVRDLLPSSAMIKDLSQEFGMPLSQEELT
DEKLFALPPQAPNLEDYHSRNSTLTLEIHAEQEKYLQWRSAMLKMKDKKHSFIQKNIT
EAYQVSKKPPKSAKVIKISAPANNAVYNYSTQTMNSTELAKKELYQEIAKEPRKRFTYS
QDYLAMVEPLDLKEEEKKAQKKSQAWLTAARGFQVTGLQSDTESSFQDLKLPIKELNE
EWKENSIFANVLEPVLDRDRWSWDRHHVDFDLYKKPPPFLELLPSPAPKPVTVRKKKGNS

PIS

SEQ.ID.NO. 313 HG1015028PIPLT00014333C16.contig.b
MAAARPCVRLGWAHLVVVVAEPNADGAAGPRQHGRHAIQVDEDVGHPLQYQLLLHDGLGS
QRLQVGTHYLDQSPTWSLPGHLCLLS

SEQ.ID.NO. 314 HG1015029PIPLT00014333C21.contig.a
MLSCTPVNSMFSGVILRGDSGPAGVPGHGLGGPALGAAGRPCQPRAVRDLAPGPAAT
EGVLGPLAVPAHRRCARFLAGP

SEQ.ID.NO. 315 HG1015030PIPLT00014333C21.contig.b
MPEPPHSMGSCAARASPTSTTPCSTAPSPIDHPRAEEYERTAPDWQAAPPAAPVRDPLG
EASWAPESGGDIESLYV

SEQ.ID.NO. 316 HG1015031PIPLT00014333C24.contig.a
MYICTHKYAYVCTHMCA YAYVCTHMCA YAYVCTRAHMRMCTHVHMC MRAYVCICKYACTH
MYIYVVRMHVHMFVCIHTKNPYIHSKMKRPMDS

SEQ.ID.NO. 317 HG1015032PIPLT00014333C24.contig.b
MHICIYVHTNMHCAHICAHMRMCAHICAHMRMCAHVRICVCAHMCCTCVCHMCA YVSMH
VHICTYMYVCMYICLCVYTQKTHTYIQK

SEQ.ID.NO. 318 HG1015033PIPLT00014333D07.contig.a

MRLGGGREKLACGRAGRAARTGRGEGRQAEAAAAGRRRRRPGTRGIPRAAGGLSRAG
TRVLRYPSSHRRQ

SEQ.ID.NO. 319 HG1015034PIPLT00014333D07.contig.b
MHRGRAPPSDSRGGRAPSVSFALEPPARVGFSGIPTVRDYSTSRLLITFFSSSSAKASLK
KNITNFF

SEQ.ID.NO. 320 HG1015035PIPLT00014333D15.contig.a
MENGYKLVGFTREGSGWAFLESHQLAVSLGIFGGLVTFREDSTNLLSKESKPACRFCQ
LTGD

SEQ.ID.NO. 321 HG1015036PIPLT00014333D15.contig.b
MIVIWKLDPWSTISFYYYFFLFCSSLWQIFSTLFFNPSFEFFTSYIVFSISKSLFCLF
VL

SEQ.ID.NO. 322 HG1015037PIPLT00014333E01.contig.a
MMPLFLIKRIKQHKCMLKELEGNKLIYWWSGSRGEILVILIFPFILLVFSKYFSNKPVL
ILFISKCFYLEKF

SEQ.ID.NO. 323 HG1015038PIPLT00014333E01.contig.b
MVGYKDFMCLFIYLLFVHLYLLIFLFFYFYFLRQDLTSLRLEYSGTIMAHSLDLLGSG
DPPTSAS

SEQ.ID.NO. 324 HG1015039PIPLT00014333E04.contig.a
MDTYDMEQFLNCLSHRGNCILYTHKLGFSERARMEHNFLTSLVSFLLVQNGQ

SEQ.ID.NO. 325 HG1015040PIPLT00014333E05.contig.a
MIKGSIQQEVITIVNIYAPNTRAPRYYNKQLITSKLIIKIGAEINEMEMKKQYKRSMKQK
VGFLKR

SEQ.ID.NO. 326 HG1015041PIPLT00014333E05.contig.b
MAGRQRRKRPPKSNQKTNNKIAGVSPYLSIITLNVSELNSPIKRHRMAEWIKHKTQ

SEQ.ID.NO. 327 HG1015042PIPLT00014333E14.contig.a
MKAMVGTIQDPTRA YSHGWSEKNGCPWHWGREGIVQQAKECLSFHEAFWWGHSVTCCLM
ILGMDSVTGWQTQFFFAVSDRLPHTYHICTLWSEQEHS�KRKVRRRKKG

SEQ.ID.NO. 328 HG1015043PIPLT00014333E14.contig.b
MVCKWQNWGEKPSGCLSLPSIVPQRWPLLSSPDPWVQVIATGCQTLQVLCWTPGSQVT
G

SEQ.ID.NO. 329 HG1015086PIPLT00014333E15.contig.a
MTKCIISLCLYDCLIQAFLLPMGWIEKYSFFFFFELESSVAQAGVQWCNLSAHCKFRL
PGSRHSPASASRVAGTTGTRHHARLTLFLYFK

SEQ.ID.NO. 330 HG1015087PIPLT00014333E15.contig.b
MYYLHDSFPFPRENCQESL TEWFLMENLDWRILJFPPTPHSVSLYWEESVETFVGDLC CF
GVYLLLLSGVSDCGKNM

SEQ.ID.NO. 331 HG1015044PIPLT00014333E24.contig.b
MATKRWSGPVKAKADQSKARVMATVFWDAQGTLFADLLEGQKTITSAYYESVLRKLARAL
VEKCLGKLHQRVLLYHDNAPAHSSHQTRGNL

SEQ.ID.NO. 332 HG1015045PIPLT00014333F07.contig.a
MPIGFESLEKIDKFLAKYKLSKPTQEKIEILYSHKVLNIQFFLETQIYAFYP

SEQ.ID.NO. 333 HG1015046PIPLT00014333G01.contig.a
MSPYECQTGISNLMCPPKILIFTPKPAPLSVLCILVNVGTGTHSVVQAKILRAVFDLSHPF
TFKSFSKAC

SEQ.ID.NO. 334 HG1015047PIPLT00014333G02.contig.a
MHPWPDVLAHTCIPSA LGSQRQADHLSPGFRDPAWATWQNPISTKKISCMWWCALLVLAT
WEAEVGSLEPGSLNLK

SEQ.ID.NO. 335 HG1015048PIPLT00014333G02.contig.b
MKTAEVKKKKAHTGRRGTGCEYLEEFCSSFELFDWFCPFKGHLJLVLFVSR LSDSSY

SEQ.ID.NO. 336 HG1015088PIPLT00014333G09.contig.a
MRLRWIQGAGAQIHQDQGRFLRARCPSGPEENPFSACLRWRTQWPPPPALRPPQGPDPPEL
DDRTPSHLSPQNPCQAPAGRGARPPASGKHSPAGACAGATLPGPLWLRRLALALARGAA
TGSAF

SEQ.ID.NO. 337 HG1015089PIPLT00014333G09.contig.b
MSFTTLPSENAGVQWCDLGLQPRTPWFKCFSSASRVAGITGTCHHTELJYCISRDRV
SPC

SEQ.ID.NO. 338 HG1015049PIPLT00014333H11.contig.a
MKSHKSIFSRDCEVHCTQPALFYISFFYFNHLGTGQKSWSEGHICIALMMGAENKQEELG
NENCFTSYILTRKATQMGITGIVQWQKDTILKNRS

SEQ.ID.NO. 339 HG1015050PIPLT00014333H15.contig.a
MPCKEPCDHQHIRGQESFITGSLMLLFIAKLTSFFPLPEAGNHSMVPIFYNFVALRMLYK
WNHTLCNLSRLAFSAKRNSLETHPGWQACP

SEQ.ID.NO. 340 HG1015051PIPLT00014333H15.contig.b
MSTSLAVGAGAVVTAESLEPSRPSVEEGQKGTAEGVCVCVCVCVCACVACVCFVF

SEQ.ID.NO. 341 HG1015052PIPLT00014333I18.contig.a
MLWDLTVLDTCMVCCGLCLACGSLLWPPPLMYKPRLCASLGRQAAVHSLTREV

SEQ.ID.NO. 342 HG1015053PIPLT00014333I18.contig.b
MVSVERGWNCLLVYLWEAAALEEEKMEKVSQYQVVRDNTFCEVVLFFHF

SEQ.ID.NO. 343 HG1015054PIPLT00014333I22.contig.a
MEYRHVAQVGLKLRLVSSDPPPPPKILGLQACATMLGHKFISGVEVLLCKYYKVEETI
NHF AKMKGWG

SEQ.ID.NO. 344 HG1015055PIPLT00014333I22.contig.b
MRFQQFLFAFFIFIMSLLLISGQRPVNLTMRRLKLRKHNCLQRRCMPLHSRVPFP

SEQ.ID.NO. 345 HG1015056PIPLT00014333J01.contig.a
MSGRVPLAEKALSEG YARLR YRDTSLIWQQQKLESVPPGT YLSRSRSMWYSQY GNEA
ILVRDKNKLEVSRDTGQSKFCTIM

SEQ.ID.NO. 346 HG1015057PIPLT00014333J01.contig.b
MRAGRATGAPSSRLRGPPCFRLALDPSLQAQWSRPWWQNHPPGCDASSLSIPHSCTRKS
CGRAPL

SEQ.ID.NO. 347 HG1015058PIPLT00014333J13.contig.a

MALSLASSPTPCFFILRKMESSFNSVSPFQVKLGHNHKFTPTYSFLSTRILLTLSTLL
CLFLLLLYKYLLSFIVSLFPFSDSPYPTKCSRLPQTTAHLMQPTTV

SEQ.ID.NO. 348 HG1015059P|PLT00014333J13.contig.b
MSLGSFGISNEQFVRHTINHAIPLTVELFNCGFVCAREIFMISTKALKAAQTALLSIFGQ
LPWLYHSLQALHPASSFLEKWSYQVLTQFLLSK

SEQ.ID.NO. 349 HG1015060P|PLT00014333J15.contig.a
MEPQSVTFIPVPANFFIKEIYFEINVIISPSFLNISKNKDDQLTPKYQSRIKVITTTVS
YCH

SEQ.ID.NO. 350 HG1015061P|PLT00014333J15.contig.b
MHSDGTLLYFHISHLSNSPSHLSYSFLITNITILAFIMAGNWWQTAMRDYHISKPTTTTFG
VF

SEQ.ID.NO. 351 HG1015062P|PLT00014333J17.contig.a
MGKDLNRCFSQEDIQMASKYKKRCSTSFIIRETQIKTTMRYVTPTRITTIKKTITNVSE
NVKELELSFIAGMNNENWHNCLGKQFSSS

SEQ.ID.NO. 352 HG1015063P|PLT00014333J23.contig.a
MRSFCDSSSLRRTTSETKSSSAQGWGPADFKGRTGAWLVGEFPNVDLNCSVFREHNQHLS
LGDVVFV

SEQ.ID.NO. 353 HG1015064P|PLT00014333J23.contig.b
MCQNMCKLSTGSPFLPMELLANEGCGSAAGLSMLKLPHSTSTPGLSLSPHSSISGSQL

SEQ.ID.NO. 354 HG1015065P|PLT00014333K04.contig.a
MPQPPSALPSLAKDTKEETEVLPLPPMNWEEDKRYTTVMGPCLRQAALLEGELLACPVMQ
DQQGNWVHEPTLNTSKEIRKSIRENGATSPFTRGLIEAIADNYHMTTPWDWSVLAKTTLG
GESIPPLEGRI

SEQ.ID.NO. 355 HG1015066P|PLT00014333K04.contig.b
MNKDP TTLAQVLFTLNFLNDNKFQSAIEKHFAKTSQDTKPSVLWKDVNSNLWCSPHDL
LTWGRGYACVHIPSGPLIPVQCIKPYHGMAGTQCSTGNEECEPVGPAAPDNAASSDNTG
PGWGM

SEQ.ID.NO. 356 HG1015067P|PLT00014333K08.contig.a
MSTICGFHSYLSMVNIDHHHYLLSCYHIPATPCMLFIIFSTLSIFAICYVISQTRKLNAE
MFYFLPNIT

SEQ.ID.NO. 357 HG1015068P|PLT00014333K08.contig.b
MLVGQAGLELLTSSDPPTSASQSAGITGMSHHARPKKNFFFSISQTLWHVPVVAATMGAE
VGG

SEQ.ID.NO. 358 HG1015069PIPLT00014333L13.contig.b
MHCQGHQKDDSPITKGNQVADKAAKQATQETYLLGTLHLNLSEFKLHYTE

SEQ.ID.NO. 359 HG1015070PIPLT00014333M01.contig.a
MPPRFKPFSCLSLLSGWDYRRALPRPANFCIFSRDRVSPCWGWSRTPDLRQSARLGLPK
CWDYRCSSRPASIFRNFKGRKNTSLTLPFFSVLATWPCQIFLLILMQ

SEQ.ID.NO. 360 HG1015071PIPLT00014333M01.contig.b
MRHLKATTLMQVTDRECYQSEKEKLEDYFGFFKNSKDYIILLYFAFESKVGIIITYNCFY
LJIEFNH

SEQ.ID.NO. 361 HG1015072PIPLT00014333M02.contig.a
MFQEDVGKYGVHLAFQKSHILCKIGMLIHGYFGLVEAILLFQKTDCLKILTVHLGDKKF
LGLLNRARKREAFSLDHDNSSSTLHYKGAISPFTLQGMKFQ

SEQ.ID.NO. 362 HG1015073PIPLT00014333M02.contig.b
METGTGCWGLGGAIEEPVAVPCSGKRTRCVASQNDEFHFSRRLRGILPK

SEQ.ID.NO. 363 HG1015074PIPLT00014333M07.contig.a
MTCMDLRNVGRSFILIKFFLTQWFTLIYQFLTLLSREKIAELTCSILTDSCPSASGKLL
KEVGVAQKYS

SEQ.ID.NO. 364 HG1015075PIPLT00014333M07.contig.b
MTLLGLCQIILGQGGWFTQYTLPMSERQLSQEGRWILGVGNRCRIPTFYSLVEVMFRT

SEQ.ID.NO. 365 HG1015076PIPLT00014333M15.contig.a
MANKLMKTCLTSLIIREMQIKAKMRYHVTLIRCYFLKITQQQKIRNVGEDVKKLKPLWTV
VGDVKWRSCWKGKYGSSQN

SEQ.ID.NO. 366 HG1015077PIPLT00014333M15.contig.b
MFTAEPKGGSNPNVHQWMNTGNKSCVWRPLRFTQLSTRVRAQPSRLFPMLHLQS

SEQ.ID.NO. 367 HG1015078PIPLT00014333N05.contig.a
MLGLSLQPQAHLSRRETLHFACPTNISNMSPNELIHSEAAWLNEWCHCLPRSLERNVGAS
VHDHAQCSPTLV

SEQ.ID.NO. 368 HG1015079PIPLT00014333N05.contig.b
MTMTMPKAPCCPWPKLWMSLSETAYSQVSLCSPCPPLTHCPHRRSNRKPDPILFKLLRV
SHCPCWNPES

SEQ.ID.NO. 369 HG1015080PIPLT00014333N11.contig.a
MAKSRFTSSWKEAGQKFRVVGPPWPGWMCYNMNHKPPSLILADRKLELWRKVSHQALRPN
KYSHSLQSYRTFTCTPFIRLPKDAWNVLAGATISTC

SEQ.ID.NO. 370 HG101508PIPLT00014333N11.contig.b
MITAHCSLTLDSSNPPHSASQELRTGSIDYKESRMSADQATRRRFPDFHYARGLHNCVC
EVMKNVLAL

SEQ.ID.NO. 371 HG1015082PIPLT00014333O03.contig.a
MLSLPQAPTRNRPCVMFPALCPGVLFVQFPPMSENMKYLVFCPCDSLRLRMVSSFIHVP
AKDMNSSFFMAA

SEQ.ID.NO. 372 HG1015083PIPLT00014333O03.contig.b
MSFDLGKNIVGHPESTCHQPRRACYKHTRHAKDGNIGPNLATTNPPHAYFSNYK

SEQ.ID.NO. 373 HG1015084PIPLT00014333O10.contig.a
MLCSCIYHSGCWRRELPRGIADFKLQHSSTVRGDVSTQNLTVITAGYKLICLME

SEQ.ID.NO. 374 HG1015085PIPLT00014333O17.contig.a
MFGPGAVAHACNPSTLGGRGSHLSTTHEVSGAPPSAQEERHMTQA YANQHFTFLMVGSKL
CIWVHLFPET

Appendix C

SEQ.ID.NO. 375 HG1014903N0 PLT00014330A02.contig

AAAAACAGCAAAAAGTGGATAATCTATGACATAATCCAATTTGAAGCTAAGTATAATAGAA
TGTTGCTTTTAGCCTCCATACTTGGGCTGCATATGGGTTTTGTACTTTTCCCTTAAGGT
TCTGCAATCCATTGCTGAAAAATTTTAATAATTTTGCAATTTGAGGTGGTTTCCCG
GATGTAGGAAATGTTAGTGCTAAAATCTGGACAATGCCATACAAATCCTCACCCTAAGACTG
GTAGATAGCTTTAGATATAATTGGCACTCAACATACCTAAGATGTAATTTTCTTCCCGT
ATCTCTGCTTAAAAATATAATTTTCTGTCTTGCCATAGTGCACATAAGTTGATCAAAACCAGAA
ACATGGGAATCAATTAATGTTAATGTTCTTTTCTCTAAACCCCTTTTCCAGAAATTAATGAAG
TGCTATGAAATTTTAAATTTCTTCTTTAATTTTTCATTTTATTTTCTGAACATGGACTGC
CTTCCCATGTTCCAAAAAATTACAAATATAAATAGGTTTACAGAAAAAAGCCTATTTTCT
TTCCACCGTGTCTCCAGTCTTTCCAGTTTTTGGCCCAAGCTAAAGCAGTAATTACTATTGT
TTGTTTTATATGCATCTTTATGGACATACATGCAAAATACAGACTTTTAAAAATGTTGTCCC
ACTTTGCAATAATCACTGCTATTTATAAAAGATAAAGAATGGCACTATATAGCAGTAAGA
GATCTTTAGCAATTCTAAATTTTATCTGGGCAAAATATCGCCAGCCTATTGTCTCTAAAGTT
AGGAAAAATTACCTTGAATTAGGTCCTGGTTCTGCTCTTTGGATATGGCTTCCAGTTTTT
GTGCACCATGGATCTTGGCTTTATTCTCTTGAAGGATATTTCATTTCCATTTCTTCTTCTT
CTTCTTCTTCTTTTATTTTATTTTGTTTTTTTTGGAGACAGACTTCTGCTCTGTGTGCC
CAGGCTGGCATGCAGTAGCACAAATCTCCACTTACTGCAACTTCCACCTGCCAGGAGAATC
ACTTGAACCCAGGAGGTGGAGTTGCAATGAGCTGAGATTGCACCATTTGCACCTCCAGCCT
GGGTGACAGAGTAAGACCCCATCTCAAAAAAAATCTCAAAAAAAAGAAAGTCAAGA

AACAACCTGCTGGTAAGGCTGTGGATAAATAGGAACACATTTACACTGTTGGTGGGAATG
TAAATTAGTTCAACGATTGTGGAAGACACTGTGGTGATTCTCTCAAAGATCTAGAACCCAGA
AATACCATTTAACCAGCAATTCCATTATAGGGTATATAATCAAAAGAATATTAAATCAT
TCTATTATAAAGATACATGTCATGTGCAATGTTTCATTGCAGCACTATTACACACAGCAAAGA
CATAGACTCAACCCAAAATGCCCATCAATGATGCATTGGATGAAGAAAATATGGTATATTT
ACACCATGGAATACTATGCAGCCGTAAGGGAACAAGATAATGTTCTTTGCAAGGACGT
GGGTGGAGCTGGAAAGCCATTATCCTCAGCAAACTAACACAGGAACAGAAAACCAAAACACC
ACATGTTCTCACTTATAAGTGGGAGCCCAACAATGAGAAACACATAGACACAGGGAGGGGA
ACAACACACACACTTGGGTCTGTCATGGGGGTGGAGGGAGAGCATTAGGATAAATGGC
TAATACATACAGGGCATAAATACCTAGGTGATGGGTGGTAGGTGCAGCAAACCACTATGG
CACATGTTTACCTATGTAAGGAGCCTGTATGTCCTGCACATGTATGCTAGAACTTAAAAA
TAAATTACATTTTTTTTTTTTAAAAAAGAGGCCACTAGAAAGGTGGCCCCCTAGAAAGCTAG
AAAGTCAAGAAAACCAATTCTCATCCTGATGTTCTAGAAGGGAACCTCAGCTCTTCTGCT
ACCTTTATTTTTAGCCTTGACCCCTTGGTTTTAATCAAGTGAAATCCATTTCAAAACCTCT
GGACCCCAAGAACTGTAAGGGAATAAATTTCATGTTGTTTTAAGCCACTACATTTGTGTGA
TTTGTTACAGTAGCACAGGAGACTAATACAATGAGCAAAGTCCAGTTATTATTTGGAAA
AGATGCCAGCATTTGTGTTTTTCGGTAGTTTCATAAAAATGTTACGAGAATCTAGAAAG
ATTATATGAAACTCAATTTTTTGATACCAGAGCCACTAGAAAACCAAGGTGAGCATTTTTTC
ATTATTGAGATACCTCTTGATATAGTGCAAAAAGCATGATTGTTAACAAAGGTTTGACTAA
ATCAGGTCTACAATGCAATAAGCCAACACATCTATTACCAACATTTTCTTATTGCTTTC
CTACATGGTTGTAACTCTGAATTGGTAAATGTAATTTTATTTCATTTTTTATTAAAAAATA
GTCAAGTTGAGTGTAGCATGCATATTTGTTCTTAAGGTATACCCATAAACTACTTCATTG

GCTATTTTCAAAAATATTGGTGCCCTTTGATGGAATAATAATAAACAACCTTTGGCT
AATTAGAAAACCTTGCCTGTTTCAGCCAGGACTTATAACATTTAATCTCCCTATATGCTT
TCACATCCCCTGTTCTACCATGCCCTAATCCCAATTTTCTCTCTGCCTATTGTGTAGGATG
CTTTGTTGTGGTTACTTGCCTCTCTAAACCAAGTTACATGTATCTTTCCATATGTCAATTAA
AATAATACGGTTGCTTGGGTGATGCTAGCATCTGTTTGTAAATCATTTCTCATTTTCATCA
TTTGAACTCTTAGAAAACATAGCCACAATAATTCAGAAATGTTAAACTAGCATTACCTCAAG
ACAAAGCACATAGTAAATCCACAGGCAATGTCTAAACGATTGTTAATGCTTATGATTAG
AATGCACTTAGTTGAATACATAAATCACATCACTCACAAATGATGCAAGAAATGGATTTTCC
ATTTCTGTGAATGGCAGATTATGGTT

SEQ.ID.NO. 376 HG1014905N0 PLT00014330A08.contig

TCAGAAAGCAAATCGATCAGATGAGCTGCAACTGTGGCTCTGCCACAGACCATGAGCTAAT
GCAGGCTGTACCCAGTGAGCCAGTGCCCGGGCACTGGAGCCGGCAGTGGGAGAGGGTGA
GCCGACATAGCTAATGCCTAACGCCAGGAGGAGAGACAAAGGGGCCCTCTTCTTTAACAC
TGTGAATAGCCCCAGAGTGAAATACAGCCTTCGACAGCTGTGGTTTCATTCTCATCACCGC
TGGGCGAGTTGGGGGTGCTAATTGTGAAGCTGGAGAGCTGTTAAATCAGACCAGATTTG
TTAAAGGAGAGGAGTAAGATGATACTTCAGACCAGTATCTGCAGTGGCACAGGCACAAA
GCCCTATCTCAGGAGAAAGGAATGAAAGGAAGGCAAAAAGGCCGTGAGAAAATGAAAGAAG
CTGGTTGATTGACCCCTCCTCTGTGTGCGCCTCAGCAACTTGTCCATATCATTTGTACAGC
ACTTACCACACAGTTGTAAAGACGGTACTGCAAGCTTACTTTCCAACGTGGCTGCATCGTT
TCACATTCCCACCAGCAGTATGAGGTTTCCAATTTCTCTGCATCCTCATCAAAAACTTGTT
ATTATCTGTCTTTTTTATTTTAGACGTTTCTAGAAGGTGGAGTGGTATCTCATTTGTGGT

ATTGATTTGCAATTCCTTAAGGACTAATGCTTTTGAAACATTTTTTCATGTGCTTATTGGC
 CATTTGTATATCTTCTCTGGAGAAATGTTTATTCAAAATCCTTTGTCCATTTTAAATTGG
 ATTGTCITTTTATTGTTGCATGGTATAAGTTCTCTATATAATTCTGGACACAAAGTCTCTTA
 TTAGTTATACCAATTTGCAAAATATTTTCTCCCATCTGTGCGTTGTCTTTTAAACCTTTCTA
 ATGTGTCCATTGAAACACAAAAGTTTTTAAATTTTCACAAAGTTCAATTTATCTACTTTTT
 TCTTTTGTCACTTATTGTATCATATCTGAGAAACCATTAACCTAATCCAAGGTCAGGAAG
 GATTACTCCTATGTTCTCTTTTGAGAGTTTTATTGTTTTTACTCTTATATTTAAGTTGAA
 TATCCATTTTGAGTTAGTTAATTTTGTATATAGTATGACATAGGATCGAGAGGATCTTG
 TTAATCACTTTTCCGTGATATACATCAGCTTCAAATGTGAAGCTATAAATCTCCCTGTAG
 GTTCCCTGAAAGCCCTCATTATATAAAGCAACTTCTGTGACAAATTTGACTTGATTAAAT
 AATTAAATTAGAGACACAGTCTTGCTCTGTCTCCCCAGACTGGAGTGCAGTGGTGCCATCTT
 GGATCACAGCAACCTTCAACCTGGGGGGGAGGTTGCAGTGAGCTGAGATTGCGCCACTT
 CACTTCAGCCTGAGTGACAGAGTGAAATTCATCTTAAAGAAAAAATAAAGGA
 AAGAAAGCAGCAAGAGAAATTTTGGCAAGATTATGAGCCAAATTTAAGAGCTTTTATT
 ACATTTCAAATGTGATTTTAAATTCTCAAAGTTGGACAGATAGCAGGGAAGAATGGTCA
 CAGACTTTTCCCCTAATAATTTAAGTTGCTTCAAACCAAGTTATTAATCTTATATGTATTT
 TTTAAAGACTTATCCACAGGGTTTGATTAATTTACACAGACTTCTTCCAAAACGTGTAT
 CTAAACCACATTTGTTGATATCTAAGAAACTCTATGAGCTTCTTCAAAGCCCTCTTAAAA
 ATACAACAAAATCATGTTAGCAACTCCAGATTCCAAAGTTGAAATCCTCCCGCTTTAGTAT
 AAGAACCATCAGTTTTTGAATGAGTCTGGATACGAAACCTATTTTCCCAATTTTGAGGAGA
 AAATCCCTTTTLAGAGACTCTTCTTGAAATGGATATTGTATTAAACATATTTGGGCTTAG
 GGCTAAAACCTTTGGTTTGTAAAGCTGAGGTTTTCATTTGCACTTACCATCAACAGCGTAT

AAGTTGAAAAGAAATCTCAGGTCTGTTGCTGCTGAGCTACTGCAAAGAAGAAACCACT
 GAGAATAAGAGCATTTCAATTATTGTACATCACATTTTGCTATTGCTGCCGTTATTCTTC
 TATTAAAGGAGGATAATCAACTGATATGATATTGTAGATCTCAGTGCCTAGTGCAATTGTA
 GGTCCCTCAATATATGTTGATTCTTCTTTTACATAATTTTGTGAGGTACAGAAAAGAAAGT
 TCTTAAAATGGTTTTAACTGTAATCATTTTGGAAAAATTAGCTATGAAGATGGAAATTTG
 TGAACAAAGGAGAACTTGCAATGCCATTTTACTGAGCATGTTACTGAATGGTTATATAA
 GGAGTAAATATTCTCTCAGCCTTAGTAAGTAAATATGAACCAATAAGAACACACAGGTGAGG
 ACTCTTTTCCTAACCCAGAAATAGATGGCTAAGGTTTACAGGGCAATAAATTAGACTATTTG
 TATCCAAAGCAATATTGTTGTCACCACCAATTGTAATAAGCATTTTATAAATTTATGCCT
 TTAATGTGTTTATTATGAAAAGCTTCTAAGTATATTGCACAAAGTGGTAATCTGCTCCAG
 AACTGGAAATTTGGGAGACTTGGTTTTTGGACTCAGCCTATTCACTATTATTGTTGTTGAT
 AGGTTTATTATTATTATTATTATTTTAAAGGCCGTGAGACAGTAGAGACTTTCTTAA
 TAAAT

SEQ.ID.NO. 377 HG1014907N0 PLT00014330A17.contig
 CATATGGGAGGCCACGTAATCACTTATTACAGTGGTTACATAATAACACTGGCTCACTGCA
 GACTCTCTTGTTTTTTGATACAGTTTCGTGCTGGCTTCATTTGCCAATTGTGTTGTTAG
 TTCGGAAGTAAGAGGGTCTTGAGATTGAGGGGTAGGGAGGGCTACACTGACTGATCCGTG
 GCTTAAGACAGGAGATTATCTCTGTACTCCAGTGGCATCTCCTTAGCCAAGATGTGAAAT
 TAAATCATAGTTCGCCCTCATTTAAAATTCTAATAAAGCACTCAAACCTTGAAAAGCTT
 TTAATTTCCCTCTACTAAAAGAAATGTATGTACCTCATAGCCCTGTGTCATTTAGTGT
 TCAGCACTTTTGGGAACATCAGTTGGTGAACTTTAAATTTGCTGTCTACTCACTGGGCA

CGGTGGCTCACACCTGTAAATCCCAGCACCTTTGGGAGGCTGAGGCAGGTGGATCACCTGAG
 GTCAGGAGTTTGAGACCAGCCTGACCAACATGTTGTAACCCCGTCTCTACTAAAAATGCA
 GAAATTAGGTGGCGCCTGTAAATCCCAGCTACTTGGGAGGCTGAGGCGAGATAATCGCTT
 GAACCTGGGAGGCAGAGGTTGCAGTGAGCCGAGATTGCACCACTGTCTGCCAGCCTGGGT
 GATAAGAGTGAAACTCCATCTCAAAAAAATAATTTTTTTTTTTTGCTA
 TATACTCTTAGTTTACATCCTCTCCCCCAATCTTACACACAGAGCCAAAGAGCAGTTAT
 ATTCTTGGGATACCATGATACCTCTGGGAAAGGAATTATATTTCCCAAGGAGAGGTCCCAT
 TGGATAAGAAACAATACCATGCGACTTGAGTCTGAATGCCAACTTACTATTACACACATCCT
 TCTTAAAAACGAACACACTGTTTCTTTTCTGACCTAGAGCATTTCTAAAGTTCTGTTCAA
 ATAAATAAAGGGCAAAATAAAGTAGATTTAACCAAGTGCCAGGTGGAATTCAAAAAACAC
 ATTCTTCCAGATAAATTCTACCTTTATGGTATGGATTTGAAAGTACTTTGCAGGAAACA
 GTCAGTACTCTTTAAAAAGGGACTGCAGGGCTGGGATTACAGGCGTGATCCACCACACCT
 GGCCCTTGCAATCTTCTACTTTAAGGTTTGCAGAGATAAACCAATAAATCCACACCGTAC
 ATCTGCAATATGAATTCAAGAAAGGAAATAGTACCTTCAATACTTAAAAATAGTCTTCCA
 CAAAAAATACTTTATTCTGATCTATACAAAATTTTCAGAAAGTTATTTTCTTATCATTG
 CTAAACTGATGACTTACCATGGGATGGGGTCCAGTCCCATGACCTTGGGGTACAAATTGTA
 AACCTAGAGTTTATCAACTTTGGTGAACAGTTTTTGGCATAATAGTCAATTTCTACTTCT
 GGAAGTCACTCTCATTTCCACTGTTGGTATTATATAATTCAAGGAGAAATATGATAAAACACT
 GCCCTCTTGTGGTGCAATTGAAAGAAGAGATGAGAAATGATGAAAAAGGTTCCTGAAAAAT
 GGGAGACAGCCTCTTACTTGCCCAAGAAAAATGAAAGGATTGGACCGAGCTGGAAAAACCTCC
 TTTACCAGATGCTGACTGGCACTGGTGGTTTTTGCTCTCGACAGTATCCACAATAGCTGA
 CGGCTGGGTGTTTCAGTTTGAAAAATATTTTGTGTGCTTCATCTTCACTGCAATTTTGTGT

AAATTTCTCAAAGATCTGAATTAA

SEQ.ID.NO. 378 HG1014908N0 PLT00014330A20.contig
 GTCGCCGGCTGGCCCGCCGGTGGCCAGGAGGGAGAGACAAGGAGGAGGAGGAGGGA
 GGTGACACTGATGATAAAAAATCCATCCAAAGAAITTTATATGTGACATCTGCAGCAGGTTCA
 AATCTCTCAGCTATGAGACAAACTGGATACTTCCTAAGGATTGCCATCTATTATGCACAG
 AGAGCCAAAATCCGTTACTGAAAATTGTCTGGGTGAGCCGCTGCTGGGTCCAGTGATGCA
 CTGGTAGACAAAGGAGTTGAAAAAGAAAGTGTGGCATAATTCTTTGGATGACTGGCAGAGTG
 TGACACCAGTATGGCACCAAGCCCTACTATGATTCCAAGGACTTCAGGACTAAAAATATA
 TATATATATAAATGAAAAAAATTTTGGTGGTATTGGTTGGTGGTTGAAGTAAATTTG
 AGCTTTACTTGGGGAATGTGATTATATTATCATTCACATAATGTAATTCCTTTGGTAATC
 AAGCTTTAAAGATCTTATATCTATGAAACTCATATAGTTATCACCTTTTAGTATTGGGTT
 ATCTCACTGTAAAGGGAATAACAGTAAAGAAAAACATCTGTTGACTAAAGGGTTAAATTC
 AGGACTATGATTTTCATGTGGAAGTGTTTGCCCTGGAGGATTCTTGAAATTTTGCCCTTCAT
 CTAGAGTATTTAAATGTTTCTATTTTGCAAAAAGCTATTATACTCAGAAATTTTCTAGGAA
 AACATCCTTTTCAGTGGACTCTTCAATAGTAAAGAGTCACACTGTATGGTGTTAGTGICC
 ATAGAACAGTTGAATTCATAAATTAATATTTTAGACCTATCAACTAAATTAATTCACCT
 TAAAATGATTTGGCAAGCAGGTATGAACTGAAAAATAAACCTATGTATTTTGTGTTTTTAG
 AAGTATTGAATCAGAAGAAAATTGGAACATCAAAAGTATTCTCTGAACCAAAATTGATGTTTT
 ATTTGGTTGAGAGTCCTTACTTTTGGTATACTATTATACCAAGTAATAGAGTATTTTGAAA
 GCAGTCTGGGAAATGAGGTGCATGGATTCTCTCAATGTTGAATCCCATTTTTCCTC
 AGTGAATCTTCAACATTTTTTAAATCTTAGCATTCACAAACAACTCTACAGCCTAGATATA

AACCATAGATTACTTTTATATAGCATATACAGTTTCATCAACAAACAGTAAAGTTTCCTT
 CATGACTTACCTTGAAATCTTTATTCAAAGGATATACTATTCAAATTTTTCAGTGTGCCAC
 TGCCACCACTGCATGTTTTTGGTCAGAGTGCAGTGCAACTAACATCGCCATCCTGCTAGG
 AAAAGATCTGCAGGTTGCTGGATAGTGGAAAGTTGCACACATCTAAATTTACAAGTCAGAT
 TAGAGATCAGAAATCTTACAATAGCTTTATGTGGAATTGATTTCATTTTCATGGTTTTTAAG
 CTGTTTTTTTCAATTGAAGCAATGCCGTGTAACCTGAAAGACAGGTGTGTGTAGCATTTTGAAT
 ACCTTAAACTCTTTTGAATCTGGAAATGAAAGGGATGTTCTGGATTCTAAAACCTCAGAGAT
 TTTGAAATTAATGCCATCATTCATTTATTGTGCCATAAATTGAACCTACCTTGAGAACTGAA
 TTTGGAAATTACTGTGATTAAATTTTGGAAAGCTTTGGCCATTTTCAATAACAAG

SEQ.ID.NO. 379 HG1014909N0 PLT00014330B02.contig
 ACTTTTTTCACCATCACTCACACACACAGAGACATACTGACACACACTGAATATTTTCTC
 CAGCATCATTTGTTAGGCTTTTAAATTAAAGCAGTGCATATAAAGGTACCTGCTCATCGTTAT
 TATTGAGGAGAGATAGCTGATCATCATTAATGGAAGGCTCTGTAGGGCATCGTCAAAACCT
 ATCCTGATCTTCTCTTTGCTTCTCCTATAGAGAAACCTTAAATTTTCTATATCTATCTG
 AGAGCAACCTGGTAGCATTTTGTAAATATTGGGTTGAGACTCTCTTAGTCTTAGATTGTGA
 GAACTTTATAATCTTGGTTCTACTTTGTCTATATGCAAGTACATTTTGAGTGTAAACAAA
 GCCCACAAATCTTTTCCCTGATTTTGTTTACATGACCTTGAAATTTATTGAGTGTAGAGAGC
 TCTGTCAGAAATCAGATTTTCAGAGCTCCTTGGCCCAAGCCAGGAAGTGGGCAATGCATGACT
 GACCAAGCCCCAGAGTGGCAGGAGGCCAACCTGCTGTGTCTTATCTTGTCTTTCCCATAGCT
 CAAGAGGACAAAGAGCAGTGTCTTAGGAGAAAGCCAAATTATTTTTCCTAACAAAGAACCATAA

AGCATTAGAAGTAATTTTCAGAAAAATGTAAGATATGTATTCCTATAAATTCAGACATTTT
AAGAGATATGACTGTAAATAAAACACATTTTACAAAGGTGCATATTACCAGCATGCTAT
AAAATGCTTCTAGAACTTATTTTAATAACATCTCCCCCTTCAGACATATAGCACATTTTAC
TCTTGAGACCACATTTTCACATCCCATCCGTTTATGCTATCTCACAACATAGAGAGGGAGGAC
AGCCTGGTCACGCTCTTGCACTTGACTGAGGCTCTGAAGAGTATGATACGCACAGCCTCT
GTGCCAGGACTGCAGTGGGGGAATTCAGATTTGTCTTTATTCAGGTCATGTTGATGAT
TAAGAAAGACCTGCCCAACACTGTCCATGCCTCAGGGACATGGGCTGATTTTATTCTTCT
GATTTTATGATGCTACTTCTAGCACCTGTCTGTAGCACAAAATCCATGATCAGTAAAATT
CCATGAAATGAATCAACTTGGCTTAACTATCACTTAGAACAAACTTTTTTTTTTTTTT
TTGAGATGGAGTCTCACTGTGTGCCCCAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCAG
CGCAAGCTCTGCCTCCTGGGTTACACACCATTTCTCCCACCTCAGCCTCCTGAGTAGCCGGG
ACCACAGGCGCCGCCACTACACCCAGCTAATTTGTGTTTGTATTTTGTGAGACGG
GTTTTCACTGTGTTAGCCAGGATGGTCTCAATCTCCTGGCCTCGTGATCTGCCTGCCTCG
GCTTCCCAAAGTGTGGGATTATAGGCATGAGCCACTGCGCCAGGCCAAGAGCAAAACATT
TTAAACCACTCTATACAAATATTCAATCCTATATAAGCCTCAAATAATTTACAGTCCCTGTG
TGAGGGTATCCTTGCAAAATGTGGGTATGTGAAACTAAAATCATGGTAAGCTATCCTGAA
GGGAAGTGTGGGCATATATTATAATTTGATGTAGTTCTTGGATTAAACCTTCACCTTCA
AGTTTGATTTCTGAGCAAAATGTCTTCCCCCAACAAATGCATAGTTGTGTTTCAACCTCCTCC
CCTTTATCTCTAGGGTATGTTTCGAACATCTCTTTTAAACGTTAATATCTGAGATTGT
AGTTGTCTTATTACAGTTTTCAGTATAAATGTCAATTGCCCTTGTCACTCTTATTCAATTG
AGAGCTCCTAAAGTCAGAGACAGAGTCCTAATGGATGAAGGCTCAATAAATGTTTTTTG
GGTAAATGGAAACAGAGGGATTACAATGGAAGAAATGAAATGAATTAGAATAATCTAAG

TGCAAAATGAATATCCAGATAAGCTACAAAGAAGAGAGATTCAAAGGAGGTGACAAATGAT
GTAAACT

SEQ.ID.NO. 380 HG1014911N0 PLT00014330B04.contig
CGTGGACGCCGGGCTCGCAGCGTGGCGGCCGAGGAGCTGAGGGAGTCGGCCGCGCTTG
CGCGGGAGTGCTTTTCGTCCTCCGGCGCGCAGCAGCGGTGTGTCCGGAACTTTTC
TGGGACTAAGTCCTGCACCCACGAGTAGAAAACGGCTCCTGGGAGCTTGGGGCTCCAAGG
GCAAGTCAGGGGACGCTGGGATGGAGGGCGCCCTCAGCCCTCGGCCGCGGTCCCTC
GCCGCCCCCTGCTCTGTGGCTCTGCCCTCGGCGGTCTGCACCTCTGGCGCGCGCAGCTC
TGAGTGCCTCTTCCCTGCGTACCGGGCGGCTCCCCAGATACGGCGCTGCTCGGCGTCG
GGAGTCCCCTCACCACTGGGGGACGGGAGCGCAGGTCCCGCCCGAAGGACAGGGCC
ATCCGGCCCCCTGGAGGCGAGGGCACTAGGCCGGGTGGCGGCGCAGCCAGCATAGCGGCC
CCCCATCCCGGCACACCGGTCTCCCGTCTGAGCGGAGGGGTAGATATGGGTCTCTGG
TTGGCGGGGGACCCCTCGTTTCGCTCTGGCCTAAGCAGCCCCCTGGGGTCCCGCGCGC
CGCAAGAGCGCCAGACTTCTTAACCTGGGCCATACCCTGTCCGCTCCACTTCCCAGTCCA
AGAAAGGAACAGCCGAAGCCAGGACGCCAAGGACCTGGCTCGAAGTCTCCAGAAATCCCG
CCAGAGCTGGGCTGGGCTCGCGGACCGCTCATGCGCGTCTCCAGAGCCTCCCCCAG
CCCAGTTTAATAGCCGGCCCTGACCAAGTCCAGGGGAGAGACCGTGCAGCCTCCCCAGCA
GACCGTGACCTCTGAGCTCCACCTTCCCTTCCCTAAGATGAGGATCCCTGAGTTG
GGGCAGAGAAAGCAGGAGGAGAAAATGGGGACCTGCGGCGCTGGGGGCCAGCTGGT
GGAGATCTCCAGCTAGGAGATCCCCAGCACAGCACTTGCCCTGCCCCTCTGTCTTGTTCAG
GGATGTAACTGCTCCTCTGCTTCTCCGGGCACCTCAGGGCACCGCCTGGGAGGGT

CTTGTA GTGCACTCAGCA TCTCTGTCCCTGGAGCGCATGGGGTGTGAAACGCCTGAT
GCCAGCCAGCTCCCCAGTCTTGAGAGGCTGGTTGAATGGCACATCTCCCTAGGAGGCTCC
CTCCCCAGGGTCCCCCTCTGCCCCCTGCTGTCCATGCAGTCGGTCTCACCCGTCAGGAAAA
CGTAGCCTGCTTGCCCTGGGTGGTTTTCCCCCATCAGGAACTGCCAGGTACTGGGCCTGGAT
GGGCTTGAAATCCCCATCTCAGTGGGAGAAAGGGGCATTGTTTAGTGCAGGGCACCTTTGA
AGCCGTCTGTAAGTCCCTGTGTCCACACCGGTCACTGGGCACTGCAGACATCAAAATGCAG
ACAGAGTTTAAAGCACAACTAGGCCTGAACCTAGGCCTGTCTGCCAACAGACATGAAGGG
GCCCCTTTTCCACCCCATCCCCAGGAGGCAAGAGGGGATACAGAAGGATAGTCCCTTGCC
TCCCCACCCAGGGTTGGACCTGAGCTGGAGTGGACAGTTGGTCAGGGTCAGAAAGAACTG
GACTTTTTATGGCCAGCTGTAGGACTTTTTATGGCTCAGCTGTAGGACTTTTTATGGCC
CAGCTGTAGGACTGTATGTCCTAGTCCCTGGTCTGGACATTGGTTCTGAGCACCCAGATGAC
CCTGCCCGATGCTGGCTAGGGAGGGTGACCCCGGAGAGGGTGCAAGAAACAGTGTCTCCTC
TCGTTAATGCTGTGGGTCCCAGGTGTCCCCGGCAAGGCCAGGCAGGAGGTGGTGACTGCC
CCATCTGTGCTAGTGCCTGTTTACAGGACTTTGAATGACACTGAAAGAGGAGGGAAGCCCC
TCCCCGCAGTCTCACCCCTGCCCAAAGTGCTGGGTGTGAGGAAATGTGGAAACATAACCCTC
GCCTGCAGTTGGGATGACCCACCATCCCTGGACCAGAACAAAGCAAAGCCTGGCAGCGGAG
AGCCAAACAGCAGGACTGACACAGGCAGGTGCTGTCTGCGATGCTCGTGCAGCCCAGGCAG
GGTGGGATCTTAGATAAAATTTTCAACAATTAAAAATAGCTGTGATCAGGCGCAGTAACCTCA
CACCTGTAATCCCAGCACTTTGGGAGGCCGAGCGGGCTGATCACGGGTCAGGAGATCG
AGACCAGCTTGGCCAACATGGTGAAACCCCTGTCTACTAAAAATACAAAAATTAGCTGGGC
ATTATGGTGCAATGCCTGTAATCCAGCTACTCGGAGGCTGAGGCAGGAGAACTGCTTGA
ACCAGGGAGTCAGAGGTTACAGTGAGCCAAAGATTGTGCCACTGCACTTCAGCCTGGT

SEQ.ID.NO. 381 HG1014913N0 PLT00014330B05.contig

AAGCGGGCCCTGTAAAGCGGGCCTGATTGGCCCTTTGAACCGGGCGTGGGATTGGCC
 GTCGCCCTGCCACCCCGGGCTGCAGCGTTCCTGGGCCTAGCCAACCGGGCAGGGC
 TGGGCAAGCGGGAGGAGCGGACCCAAAGATCGGCTACGCTGCGTCTGCTCAGGCTGCG
 AGTTCCCGCTCTGGGACTCACCTTGGGAGCTTACCAGCGGACTCTCCGCAGCCCC
 TGATGGGTGTGTGTCTAGACCCGTTTACTAAACACCAGGACTGTTGTAGGCGACTGTAAAG
 AAATAAGAATAATTGCATGTTACAAATTTCTTGAATGCTTACCTTGTGCTGTTCCTTACT
 CTACAGAAATTTTCTAATCCCGTGGCGAACCCCGTACTTCAACGTATGACATTCTGTAA
 ATCTACAAACAGATGGTCCCTGCCCTTGGAGAGTTTTTACAATTTAGTGGTACTGGA
 CACGGACCTAATTTAATTCTCACATGCCCTTCCCTTCTAATAAGCTGCAATGTTATGTATC
 CCTATTCTACAGATGTGGGGAATCCGAAATTTTTAGCTTTTACTGGGACACACAGCTAGC
 AAGCAACTAAAGCGAGAAATTCCAACCTGAGCTCCAATCCGAACTATTTTCATAAAAGGCA
 GGATCCCAGGCTGAGGGAGTTCGAGGTTTGGGAGACAGTTTAGCAGTCGCTGAGTCCAAA
 CATGTTGTTTTCCTAGAAAGAAAACTGAGATCCAGGAAAGGAAAGTGATTTTCCAAAGA
 CACACCTGTGAAAGGTGAGGAGAACAGCCCCAGTGAGGCTGAGGCCACACAATTAACCCCA
 GTCTTTGCACTACAACCTACCAAGGATCCTGGTTTGGCTCATTTTCTTCTGCAAC
 ATGATTGTTCAACCTTTTCATTACCAGCAACATACCCAAAGACCTAGCTATAGACTCGTTG
 GAATCTGTACTTAGCAGCTGTGTGAATCAAGTGTTCATAGGAATAATAAGATGTATTC
 ATCTTACCTTTTCAGAACATTATGAAGAACAAATGTGGTAAGACATATGAAAATGCTTTG
 TAAAGTGTTAGATGCTTAATACACAGAAAGACATTGCTTCCAGATCTGTTTCAACGTACCC
 CAAGAAATGGAATAAGTGGAATATGCTGTTTTCATGCACTGTGGCAAGGGGTCTCATCTT

GGCTAGACTTGAAATAAGTGGATTGTGTCCTGGCCTCTGTGGATGGCTCTTTCTCTGGGC
TTTTGTGATTTTGATTTTTTATTGTCTCCAGCAGCCGTGTTGAACCATGAAGGACCTCA
GGAATGAAACCCACACAGCAAGCAACAAAGTGAAAGAAAGCGTGGATTCTTAATGCAG
GAGGCACTGAACCAATCAATACCCACTGCCCTCCAGACCTCCACATGAAAGAAAAAAACT
GCTGTCCTGTGTAGACACTTTTATTGTTGTTTCTATTGTATGAGGTCAAACTTAATAA
CAGATATATTACACTCTAACAGATACATTACATTGTATAGTAGAAATTCTCAGACAGCAAC
CTTATAAGGTGATAGATGCAGAAACAGGCCCAAAGAGTAAAAATAATTGTTTCAGAGTTT
TCCAGCCAAAGTGGCGAAGGTGAGATTCTTACCCTGGGTTTGTCTGAAACCAATTTTCACTA
TCITTTTTCAGTTATGATACACTGCCTATTACGAATGAATCCITTTTTTTTTTTTCCA
TTTACTTCCATTGTGATTTATTCCAGGGAGTGACACTTTTCTGTAAAGGACCGATA
GTAAATATTTAGACTTTGTGAGCAATACAGTCTCTGTGTGCAACTACTCAACTTTGTAAAT
GTAAAGCAGCCATAGACAAATAGTAAATGAGTAGGTAGTATGACTATGTTTCAATAAACTTT

A

SEQ.ID.NO. 382 HG1014914N0 PLT00014330B11.contig

GCGGCGCCCGGGGCTACGCGCCGCACTGCACCGAGCGGGCGCAGCGCAAGCTTGGGTGT
GAGCCCGGAGCCGCTTTGCTTACCGTCTGCGGTGCCAGCCGTCGCTAGGAGGTCCGC
GGGCCCTGCGGCAACCCCTCGCTACAGACGCTGGGCGGGCGGCGACACCTGGCTCATGGCC
CCCGCGGCGGCTCCGTCTTGGCCGTGAGGGCCTCAAGCCCCGCGGACACCCACC
TCGTACGGCGTCTTCTGCAAGGGGCTCTCCCCGACCCCTGCTCGCCTTCTTCGAGCTGGCC
TGGCAGCTGCGCATGAATTCCCCTACTTCTACGTGCGGGGCTCGGTGATCCTCAACATC
CGATTGCAGGTACATATTTAGAGCCATGACTAAGCTAACGGCCTCCGGGCGCAGCATGAT

GGCCGACTCCAGGGTCCGTTGCGGCGGCGGAGCAGCCAATGGCGAGCCCCACAGTCT
 CGCGAGAGTGCTCAGGCGCTCTTCGTGGCTGCCCTCTTAGCTGCTAGCGGAGCTCCTCAG
 GGGCGGCGGGAGCCTACAATCCCTAGAAAAGAGAAATACGCTGTTCGGGAAACAGAACTG
 CAGTTAAGACCCCTCGAAACATCTAAGAAAGTGTGCATCCTAAACACCTGACGAAATTTT
 AGAATGTGACAAAGCGCAGAGGATGCATTATTTCAAAAACAACAGAAAGGCTAAAAATTTG
 CAGGAAAAAGAAAAATCAGTAAACCGGGAATCCTCGGACTGGATTGTAAGCAAGATTTCAA
 TGAATAAGAAAGCTGAAGGTATTAAAGGCTGTGATATAGAAGGTACATATTTTCATCCCAAA
 GAGAAAAACAATAATAATCAGAAAAATTTTCGGTGAAAAAAACGCAAACTGTACAGGAAAAAT
 CATCCTCCAAAGTACCAGACATAAAAATGCTGCAAGCTTTTGAACTAATGGCGAGAGTGTA
 GAAAATGGGCTCTACTTCAGTGATCCTGTGCGAGGACGTGGATCAAGACTTGGAAACCGCA
 GAAAACGAAATCCCATAGTAGCACAAAGCTTGGCTGTTTCAGTGAATAACATTTAAATAAT
 CGTAAAAATACAAATGTTGTTTATGGTTTTTATGTTTAAGGGCATACTTAATTATGGTTA
 CAAAGTGGAGTGCAAAATGTTATTTACCATGTTTTTAAAAATACAGCCGGAATAACAAGTG
 GGAAATGTTGAAGGAGGGCGGGGAAAGTAAATGGAAATGGGGGTTATGTCCTTATAAAGTGG
 AAACTTGAAAGGTACTGTCTGTTGTTGAGTGGGAAAGACATTTTATTACTTACAGGGT
 AGCCATAAAGTTTCTAAAACGGTAATATATTAAGAGGGAGAGTGGTAGGGGAGAACAGT
 ATGAAGTCAACAGGAAATGGCTAAGATGGAGAGCTCAGGTAGAATAGTTTAAAAAATGA
 AAATTGATAGAGTACCTGGTGTATTTGAAATACAGCCTAGAGAGCTTTCATTTTCCAAAGA
 GTGTGGGGATAAATTAGTGAATGCATATTTTCAAAATTAAGAAAAACAGAACATATAATT
 TTGAAACGCTGAGAAAAATAAAAAATTTAAAACTACACTGTATACACGGTATACCTCAACT
 GTGAATAATAGTGCATGTAAAGTCAGTCATAATTATGTAAATACTGAATACTGATTTT
 ACAACAAATGATTCTATAACGCTGGGAGGATGGGAGAAAAAGAGATGTGTGTGTGTGTAAT

GGTGTGAGCCAATCAAGATCTGATATAGAAAAATGCAGAAAAAGGAGTATACACATGTT
TTCCTCGGATTAAAAAACTAAATAAATCACCAGAGTGGTAAAGTTTTTAAATCAAGA
TCTGATATAGAAAAAGTGCAGAAAAAGTAGTATACGCATTTTTTTTTTTCAGGAAAAAATTA
ACTAATAAACCACTAAGAGGGGTAAAGTGTTAAAAATAGTTGCATTGTTCTATATTCCACT
CATTGTGTCCAGCAAACAGTGTCTATGGCTTTAAGTAGTAAGCATCTTGCCCCGTCCAA
CACCATGTCTGGCCTAGGGTAAATTGTTCTTTCTAGTCITTTATTCTTTGTCTCCAGCCAAA
CTCCATTTGGGCACATTCTCCTATAGCCTTTTCTACCTGAAGATACTCTGTGTTTTAAAG
CTAAGCTTAGGTGGCGCTTTTCCATTAAATTTTCCCTGGATTCCCACTGACCATATAGGA
GCTCACITCTTTTAATCCATAAGGCCATTTTTCATAGGTTGCCITTAATTTTCCCTAATCGT
GCATCAACTGTCTGTTTTATATACCCAAGACAGGTTTCCCTAGACTGCGATAAGCCAAAAC
ATTTTAGTCTAAAAATATCAGAAAGTGTAGTTTAATCAATGAAATAGTAATACCAAGGGATT
TAGAATCGTGGACATCACTGTTTCCAGAGCACTGATGTCCCAATTTGTAAACACAAAAAGA
CTGTCTAGTCITTAATCCTGAAATGGTGACAGAGTAGGATGCTCCCATTTTGGGTGACTATGT
GAACATATTCCTATAAATTTTTCCTCACATCAGTCATTTGTAAACCGAAGATGAAGAA
ACAGACTTTGTATTTATTACATCCTCCAAACATAATTGGCAGCTCTGTGAGGGAAGACACT
ATATCCTATTCAATTTTACACCCGGTAGTAACAGTATGTGTGTACTGTATATATACAT
ATATATATACATACACATATATAAATTTAAGTAGTAACATATATATGTATATATGTATAGT
TACTACTTACTATATGTATACAGTATATAAGTATATATGTATAGTTACTACTTACAAT
ATATGCATAAAGTTACAGCTTACATTTACCTCGTAATATATGTATATATATATATTAGGA
GTTAAATGTATTAGGAGGTAAATGTAAACTCTGTGGTTTTATCATTTATTCCACAAATTGCTC
TTAATTCITGGGGTATAGTTTGGCCTCCTACATGATGCAATTCAGTGAAATTCGCCAGGAGA
CAAAACAAGATTTTGACCTGAATAAACCACTTATTGTGTGTAATTCCACAAAGTGAGTTTATT

CTTCCTGGCACACTTCCAGGTAATATCCATTAGGCTAAGGATTTTGTGTTCACTGTTCT
TATCCCCAAGCCTAGAGCACTGCCAGTGCGGAATAGCTAATAAATATTGTTGAATGGAT
AGATGAATGGCAATTTGTGTTGTAATGCTCTACATAGTCAGCAGTCAATAAAGGA

SEQ.ID.NO. 383 HG1014915N0 PLT00014330B13.contig

AATTTTATTATGTTTGTAATTTAACTGGGAGAAAGTAACCTCTGTGATGACCTCAATGTCC
CTAAATCTGAAATTCTAAACTTTGTACAGATGTAATAATAATCTCATGGGGAGTGGGGG
AAGGGAGGGTCTTTATTTTAGGCATGGCAAAAGTTTCCCTTTCCTTACTTAAGATTTTA
TTTTAGCAAGGAAATTTTGAACAACTTCTTAACTAACCCAGGGACATGCATCTGAGCAAA
GCTTTATGTTAAAGGCAAGGAATAGCCGAATTCATGGAAGTACTTCTCTATAATATTT
TGCTTCATTTTCAGTTAGGGTAGAATGAAGAAGAAATTGATGGCAGTCTACTTATTT
TTCATTCCACTAGAAATTATTTATCAGTTGTAAATTTTCACAGCTGTGTGTTTCACATG
GAATTTTACAGGCTGACATCTTGACAGAGGTATATTGTAAATAAGATACTCTCTCCGG
GAAAAGAGAAAGGCAGTTTAGAGGACAGGTAGAGAGGAAGTATACAGATAAGGTGTGCAGG
TCTGCTCAGGGTTCGAAAGCAGTTTCCCTGGAAAACTTGAAACTGTGGCCAAAGGGGTAAG
TAAGGCAGTATATGAGGAGGCAAGATGCTGGCTATGGAAAGTCAAGCACAGCTTCACAATG
AAATCAGATTTTCTTTTCATGAAAGTAATTGAAATTAATAAAGGTATTAAATCAGCAATGT
TACCAAATAATTGGATCTTTTGAGGAAATAAAACTCTTGCTGAGAACCTACAAATCTGGC
ACTTTTATTAAAGAAATTGTGTAATGTAATTTGACAGTTAATTTGACAGCTGAAAATGTG
GCAGGACTGTTAAATAAATAATCCAAATGTAGTCTCATCCAAATTTTGACTGATTTACTA
AATTACAAACCACCAAGATATCATGTACCTTTTATGACACATTGTTTGGTTATAGAGTT
AAATTCAATGGCATCAATACTTTGTTGAAATACTTTTAAAGTAAACCCTCATTCCAATTT

TCAGGAGATAAAGTTTTATAACAATAATCTGGATGTTTATAAATTCCTGCTCCTAGGAGAA
AGTAAAAGTAAAAATAAAAAGTGAAACAATAATCAGAAAAATACTTATAACTCAGAATT
TGTGATAATTTCTGGTGGCAAGGACTAAAGTTAAAGTTGACTTATTTCCCTCAATGAACCTA
CTGATATAATTATGCAATACTTTTCTCTTGTGATTAGTGAAAGTATTATTTCTTGGTAGCACT
TTGCTCTTTAAATATTCATAAGTAACAACACTGCAATTTAAAAAATAATTTTCCATTATGA
GTACTTTTTTAATTTTCAAACTAATCTTTTTTCCCCCAAATCTGTAGTGAGTGAGAGAGTAGG
TGATGTTTAATTCGTGGTAGGGTCAATCCAATATGCTATGAATCTCAGTTCTGCTATGTAC
TTGCCAGACAAAGTTACTTGAACTCTGAGCCTCAGTCTCTCATATGTAAATGTGGCTGAG
AGCTACCCCTCACTGGGCAGCTGGAAGAAATAACATGAGTAATGTAAAGCAATCTCTTACATG
TTATGTCATCAGCATTTAACAATTGATCCTTATCTTTGAGACATTCATCATTGCATTCC
TGGATAACACACTCCTCAGCTTATCTATCTACTCTATTGGTCACCTCACTTTCAAACCTCAT
TTTCAAATCAITGATTTCTTCCCTCATCTCCCTGATTTCTTAAATATTTTACTGCCCCCAAGGC
TTAGTGTTTGAACTTCTGTCTCTATTTACATTCACTCCCTTGGTGAGTTCTTTTCAGTCT
CATGGCTTTAAAGAACAATTTGTTAATGAATCCCCAAGCTTGTATTTCCCTGGTAGGAGCTCT
CCATTGAACTCCAGTCTTATCATCTACGTGGATATTTAATAGACATCTCAAACCTGAATCT
AATATCATCTGAGCTTCCCCTCCCCAAACTGATTGCTTCCCCCATCTTAGTCACTGGGCAACTT
CAAGTTGCTCAGTCCCTAAATTTTACGGTTGTCTTTACTTTCTCTCTCACCCCAAATAT
TGGAATTGTTTGGAATCCTGTTGGTTCTACCTTCAAAATATATCTAGAAATCTGTCAATTG
CTTATGAGCCCCATCCTCCAAACCAACCATTAATTTCTTAACTAGTTTCCCTCAATTGTG
CCTTAGCACCTACAGACTATTTCACACAGAGCACCCAGAGAAATCCTTTAATCATGCTAGC
TCCCTGCTCAAAACCTCCTAATGTCTTCCCTTTTACTCATAGTCAGGACCCAGAAACTT
CAACGGCCTACAAGACTCTACACCAAGTGCGCCCCCAACCTTTTTTGGCACCAAGGACCAAGTT

TTGTGGAAGACAATTTTCCCACTGACTGGGAGGTGGGGTGGGGCTGTGTGGGGGTGTG
AGATGGT

SEQ.ID.NO. 384 HG1014917N0 PLT00014330B18.contig
CTGCTGCGCCCTGCGGAGCTCCGAACACAGTGGCAGAGGCTGGCTGTGGCAGATGCAACT
GCAGGATGACTTGA AAGTAGGGCATCCTTCACCCATCTGAAGGAGGAAATAGTGGCAGG
TGACAGTCTGCATGTGCAGTTTTTCAGATGCCTTCAACCTGAATGACATCTACCTCCATCAG
GACCCAGATGTCTGACAGCCCTGTGTGACACCAAGATAAGTAACGATCCCGTTGAAGAT
ACCACATTACACTTAGGCTCCTGTGTGGCCATGACAGCTTCTTAGACTTGTGTTTTTAATGAC
CTTGACAGTTTTGTTTTTTGTTGTTGTTTTTGAGACGGAGTCTCTGTCTATCCAGGCT
GGAGTGCAATGTAATGGCTCAATCTCAGCTCACTGCAACCTCTGCCCTCCCGGTTCAAGT
GATTCCTCGCCTCAGCCTCCCCAGTAGCTGGGATTACAGGCATGAGCCACCGTGCCCCGG
CTGCCCTCCTAAATTCTCAAAACATTCCAATTTTTATTCTGTGGTCTTACTTTATTCTGTCA
TGTAGAGAAATAGTGGTTGAGCCTGTGTCTTTTGAGAGTGGATGCTGTAGTCCTTACATG
GCTGACTTTCAGCCCGTAAGCTTATAGTAGAGGATGGGATTTTGTATCAGCTGCTAAA
AGGGAGAAAGCCAAAGCCAGAACTAAGGGTTCTGATGAGTTGAAGGAGAGGGAAGATGCCC
CTCAAGGCAAAAGAGCCTCTGAGTTAGAAATCTGTGCTTCACCAGAGTGACAATGGTAGAAC
TTCTCCTTCTCAGTATTAATTCTCCTGGAGAGAACTCTAGGGTTGTCTAAACCTGTTCAA
CTAGGCTGACTACCCAGCAAGTGATGGTTGGACATAAGGCAAAAGAACAGAGGGGTGA
AAGGGGCAGAACAGACAAGTTCTGTCCCAGCCTCTGCTACCTCTAACCCCATGGCATTC
TATCCTTTTCTACACTGGGCTTCCATTCTTACCCCAACAATGATCTGTGTTCTTCCAGGTG
CCGTCAATTTAATTTCCCAAGACACTTGACCTCCTTCTGATTTGTGTACTCCCTCCAAGGCT

GAGTTGCAGTGAGTGACAATAATCTGTGCTAATTAATTAATCTTGCCAGAAGACTCAAAGG
GTTTATGGCTTTTACTAACTGAACCTCTATGCTAGATGTTAGGGATAAATGTTAACAGGA
CACAGTTCTTGCTTATCTTGCTATGGTCTGTCAAAGCCTTATACAACTAGCCTTATCTCT
CTCACCGTCTTATCTTTATCACCCGTAGATTCCCCTTGTTGGCCACTGGTTTCTTTTCAGTC
CTTAATTAGCCCTTTTGTCACCTGTCTACACATGCTGTTTTTCCCCTTCTCATTTCCATC
TTGACATTGTCTATTTTGAGATCTCAGCTCAGTTGCAGAGAGAAGCTGCCCTTCCCCCATCTGG
CAACCCATTATATGTGTCTATAGTACCATGTTGTACCATAGAGCTAGACACAGGTGCCATG
TTGTGTCTTGAAAATGTATTCACCAGCTTCCAAAAGGTTTACCTCAATCCCCTTTACTCAAG
AAGCCTCAGTTCTACTGACAGTTGGTAATAATAACCTCTCTGTGCATAATGTACCCAAAAT
AGAGTAAGAATAATCATGCTTTTTCAGTAATACTCCAGTGAATGAGGCTAAGAGTACCATT
TTGTTCTTATAAAAGAAATTTTTTTGGACATGAATACAAAGATGTCAGGTTACCAAATCAT
TTGCTAGTAGATCCTAACAAATATCACCTATAGGAAACTGAACGTAGCCTTTAAACATTAA
GTGATGATAATGGATTTGGCCGGCGCGGTTGCCCTATAATCCCAACACTGAGAGGCTGAG
GTGGGTGGATCACTTGAGGCCAGGACAGGACCAGCATGGCCAACATGGTGAAACCCCGTCT
TCTACAAAAGAAAAAATACAAAAAATTAGCTGGGTGTTGTGGTGTATGCCGTGTAATCCCA
ACTATTTGGGTGGCTGAGGCACGAGAAATCGCTTGAACTTGGAAGGCAGAGGTTGCAGTGA
GCTGAGATCGTGCCACTGCACTCCAGCCTGGGTGACAGACTGAGACAGTCTCAAAAAAAA
AAAAAGAAAAAATATGGATTTTGCAGAGACTTGCTATTTAGATTTTCAGACATCTGTGTTAACTA
AAACACATGTGTAGGCTTTTGTTACTTAATTCAGTAATCTGTAATAATCTTTATATTGA
GAAAAATTTGTGAGACATCTTTGTGTAAATTATAAATTGAAGAACCTCTCTTACAAGCAGG
CATATTGGTAAGTAGCTGCGAGGATATAACTTATAACCAGATTGAAGTGTATAATTATAA
TATGTTATTAATCTGGGGTTCTATAA

SEQ.ID.NO. 385 HG1014919N0 PLT00014330C06.contig

TGGGCAGCGTGGACAGGCCCCACCATGCTCCGCTGACACCTGGCACACCCAGGGAGGG
 CCCCTTCCTTACCTGGGGCCAGCTGTGGGCACTTGTGGCCATGGCCTCCTCGGGGCCGGA
 CTCTCCAAAGTGGCCTCCAGGAGGGGTTGCAGGGCAGCGGGGCCCTCCGCCGTTCCCTCC
 CGCTTTTCTTGACGGGGCCTGGGGCTGTGTGGGCTTCAGACTAATCTAACTCCAG
 AAAAGAGAGAGAAAACAGTGGGATGGAGTCTGCGCGCCAGTGAGTGGAAAAACAGCCTC
 CCGGCGCCACCAAGGCAGAGGAGGCGGCCCCGAGGGCGGCGCAGAGAGAGAGGCCGA
 GAGGAAGTTGAGATGCCTGTTCTGCAGCTGCTGCGCGTCCCACCGACCCGCTGCTCGCCG
 TCCCTGGGCCGTGCTGGCCCTCCGCCCTCGGCCAGACCCGACCGGGCCATGCTCCACA
 CCAATGGGACCAGTCGGGAGGGCCACGTGGGGATGTGAGGCAATTGGGGAGATCTGACCC
 AGACCTGCCCCCCATGACAGTTTCAAGAGCACCCACGTACTTGACCCCGCTGGGCAGA
 AGCTACCAAGGAGACCCCGGAGGGTCAAGGGAGGGTGTGTCAGGCCCGCGCTGAGCC
 CTGCAGGATGCTCAGGGCCAGGGCCCCACCCTCTACTGGCCCTTCTCCCTCAGAGTTGTC
 CATTGACTGGACAAACCATCCCCAATGCTGACTCTTTTGACCTATAGTCAGCTAACTCCCC
 GGGACCTGCTTCGGATCCCTGGAGGGCTGAGGATTGAATTCAGGAGCCTAGGGATTCTCC
 GGGTCCTGCAGGCAGAGCTCCAACTGGTGGTGAAAGTTACAAGGAGGCAGGTTTCTGAA
 GGAAGCCCTGTGTAATAATCAGAGCTGCCCCAGCCTGTATCTAAAGTGTCAAGATGCGG
 GCTGAAGCGTCAACGGGTACAGGAATGCTGCACAGGAGAGTCTACGGACAGGCTAGATGACC
 CTTAGCATGGAGGGAGAGGACCCCCGACATCTGTACGCTCTGCAGTCTCGTGGTTCCGT
 CATGTTGAGCTGCCGAGTGACCCGCTGGCCTCCCTGAGCTGGGCTTGACCATATGTACG
 AAGGAAAGAAAAGGCATCCCATGGATGTCTCTCCTAGTTCCCTCACTGTGGCCCAGTGTGT

ATATCAACGTCAATGATACACCTCCTGTGGAGGTGATTAGGCCATGAGGGTGGAGC
 CCTCGTGAAATGGTGCCCTTATAGAAAGGACCCCTGGACCTGGACCTCCACTCTG
 CAAGGATACAGTGAGAAAGGCAGAACCCAGCCTCAGGTACACTCATCTCAGACTTGCAGCC
 TCCTGAACGTGTGAGAAATAAATGTTTGTGTCTGTGATGTTTGTGTTACTGTGGCCCAA
 GCTGTTGGACACAGGACACCCCTGGGCTCCCAGGCACCACTTCCCTCAGCTGCTGC
 TCTCCTGGAAACACGGTGTGCCTCACTGCCCATCCAAAGCCACATCCCGGCTCCACCTGC
 TCACTCCTTGTCTTCCACTCGATGGTCAAATTTTCTCTTACCAGAAGCTGGCCAGG
 TGCAGTGGCTCACACCTGTAAATCCAGCACTTTGGGAGGCTGAACACAGGCAGATCATTTG
 AGGTCAGGAGTTCGAGACCAGCCTGACCAACATGGTGAACCCCGTCTCTACTAAAGATA
 CAAAAAATAAAAAAATTAGCCAGGTGTGGTGTGTACCTGTAAATCCAGCTACTCG
 GGAGGCTGAGGCAGGAGAAATCGCTTGAACCCAGGAAGCGGAGGTTGCAGTGAGCTGAGAT
 CATGCCACTGCCCTGGCGACAGAGCAAGACTCCCTCTCAAAAAAATAAAAAA
 AAAGCCATGTCACAGAGCTAAGAGCAACCCAACCAAGTATCTGGGTTCAAAGTCTCAGTCT
 GGCCACACACTAGCCAGGGACCTTGGGCAAGTCATCTAACCTCTGGGTGCTGCAGCGTTC
 TCATCCGTGAAATGGGGAGAAATGATAGCGCCTGCCCTCCAGGTGTTGTGAAGATTCAAT
 CAGCTGTGCAGGTCAAGAGTTTCGCAGTGCGCCAGGCACACAAGGAGTGGTCGTGAACCA
 TCTTTACAACATTGTAATAATAAGTAGTATCATCTCTGCACCTGAGCTTCAGTAGTGCA
 TTTCAAAATCTGTATTAGTTTCTATTAACTGCTGTAACAAAATGTCCA

SEQ.ID.NO. 386 HG1014921N0 PLT00014330C12.contig
 AAACAAAGAAACAAAAAGAACCCAGTGGGGCTCTGGAAATCAAACTGGCAACGCTCG
 AGATCGTTTCTGGGGCGGAGGCTGGAAAGGGGAGCAGGCAGGAGCGGAAGCCTGAGCC

CGACCAAAGACAGCTGTGCACGCCGGCTCCCTCCCACTCCAACCTGGGCTCGCCGTCCGG
GCTACAGAGGGACCCTGCGGGGGCCCTTCGACGACCGAGTTCGAGTCCCGCCGCCACC
GCGTCTCCCCCGGACACACGGGAGCGGCCTCGCTTCCCCAGGTGGGAGCTTTTC
TCGCAACCGACTGGGGCTGCAGCTGCCAACTGCAGTCTCTGTAGGAAACAACCTTCTCC
TAAATGACACGTTTCTCCGCGACGGCCTCCTGCATCTGGAGTCGCCCTTCAAGCCTATT
ATTCAACCAAGTGGAACCTGGAGGAGGCA GTTGTGCCCTGGTTGTGGCGAGCCCCAGGG
TTTTGAAGCTGCCAGGAATGGGATCTGGGGGCTGTGTAGTTGCCCTAAGTGGGAAGTTGTG
TCACTGGGCCCGTTGGGTTCACTGTCTGTCACCTGACTGCTCTCCAGAGACCATGGAGCCT
TTCAGAAAGGGCCAGCCCTGGGGGATCTACGACCCCTTGGTTCCCTTCAAGAAAGCTCAGA
CAGCGAAGGAAGAAAGGGCAGGGCAGCCTATGAGAGAGAGAAGGCAGAGAAAAAGGAAA
GGAGATGGGAGGAAGGGAAACGCCTTCTTTAGAACTCAGGAAAAACAACATCAGAAAGTGG
GCATTTCTCCCTGTATAATCTCACAGTATGACCAGGAGAAAAAGCTGTTACAATAGAGAGC
AATTGTGAGCATCCAAATGAAAAATGTACTCAAAATAACTGGAAGAAGATGCATGGTATGAT
CAGGGTTGAGCCCTGTAAATGTATTATTAAATCATTTCTATTAGCAATATACTCATGAACCTTT
TTCTATTGCTCAGCCCTGGACATGGCCTAGCTCTTTCTACCAAAGCAAATATGACGTACT
GGTGACTGACTGAAGTAGCATATTAAATAGTTGAAAAAGCAAAACAAGCTCCTTAGCAATT
TAGAAAGTTTTTGAAAAACCCATTGTTCTCCCTAAGGACACATACTTCAGTGGCCTCTAAAT
TTTAAAATGAAAATCTTCAGGGAGATACAAATACITGTGACTTTTTTCTTTATGTCTCTAA
GACAGTTAAATTTTAAAAGGGGAATGAATTCCTAAGTAAACCTGGCCTTGAGTACAACTT
ACACTGCTCTGAAAACTGGAGGGGTATTAATTGCAAAAAGGATTGATTGTGTGTGTGTGT
GTGCATGTGTGTGTTTAGCAAATGTTTGTGTAATGATTAAAGAACATTGATGGGTATT
CGTGTCAAAATAGATGGGCACCAAGTTTGTGTTGTGCCACATAATGAGCTCTGTGACTCTG

AGCAATTTATGTAAATTCCTGTGCCTCACAGTTTCTTATCCTGAAATGGAGAGAAATAAT
AATGCCAAGTTAATTGAGCTATAAATGGTAATAATAACAAACCTGTATTAAATGGCTTACT
ATGTGCCAGACAGTATTCAAATACTCTGCTAGTAAATTACTTAGCATTTTGTCTGGAGCA
TAATAAGACCTGCATGTTAGCTTTTTAGTATTAGTATTACTACCAAGGGCTGGCAG
ATGCATCAAAGAAGTTTGCAGCCAACTGGGAGTTGTACCAAAATGCACAGAGAAACAAAGTTA
GTGAAAGACCAAGGTGATGCATCCAAAGATGTAAATGTTGGAGATTTCTGAAAGGCTGCT
AAGTAGGAAATCAAATCTAATGAGGTTAATTAGGAAAGACTTTTTTTTTTTTACAG
AGGTGAAATTTTCAGAGTTCAGAACTTAATTGGAATGGGCTTGAAAGTGAGTTCTCTAAAGA

A

SEQ.ID.NO. 387 HG1014922N0 PLT00014330C14.contig
CTCTGGCCCTTCATTTCCCTGTCCCTACACTGGTGGATGTTTGGAAAGCCTTCCTACACCAGA
GCGGCTTGCTGAACATCTACTGGATCCAGTTTATGATGGGTTCTTAAAGTTTTACTAC
ATAGTTAGGTAATAGTCAAATAAACAACAATGCAATTATCAAAACCAACCCAGTCAACCA
GGATTGAAAGTGTCAAACAGTGCACAAATCAAGGCCATTCTTTGTTCATGATCAAGAGCAC
AGAGCCTGGAGTCAGACTGACAAAGGTTTGAGATTGATGAAAGATTGTCAGGTGACTTTT
TGGTCTCTCCTACATAGTCTGGAAATCAGAAATTTGGAGCTTTAAACTTTAACATGGGTAG
GATGCCAACTATGCTTTAAAGTTATTTTCTTTGATGCAGAAAGACATAATTTTAGTATA
TCATAACATTAAATCACACCAATTAATCTTTTTTCTCTGCAATATTGAATCATAGGACTTC
TTTTCACAAACAATTTTGTGTTATAATGTTATTTTAATGTTACTAATTAGCCCTGCAGA
ACCCCTATGACATCCTTTGGCAATCGTCTCCCATCGTCTCTCTCTCTCTCTCTCTCTCTCT

CTCATTTTAAATAAACTGTTGTATAAAAATTGCATGAGGAATAAGTTCCCAACTCCCGT
GATCTGCTTATAAATAGCTGTACTTTCTTCATATAGTCTTAAGAGGCATGTCTAGTTTAAG
TACCTTCCAGAGAAGGTGTATTTATGCAAAACTTTTCAGGGGCTTGTTAAAAAGTGAAG
CTTTTGCAATGATTTGTATAAGAAAGTAATTAGGATATCTATCTAAATGATAATAAGAAATT
CAATGAGTACAGATTGGCAATTCCACTTTTGAAGATTATGGAAAGATAGAGATTTTATATT
TTAAAACAAATTTGTTAGGAATATCAAAAGCAGAAATAATTCTAGATTTGGATGAACAACCA
TATAATTTTGTTTAAAAAGCAAGAAATTGCCCCAGATTACTTCTTGCAATTAGCAATGTATA
TATCAAAAGACTCCGGAAAAGGAATTTTCTTTTAAACCACTAACCACTCCTTGCAAAA
ATAAGGAGCATTAGGACTTTCCTCCCTACCCCTTACTTTTTTAATCTACTAATTGTTCA
CACAAATTATAAAATATGTAGAAGATTGTTTTATTATGTTTTTATCTGTTTGTTCAGTGAT
AGTGTTATATAGTAAAAGAACACTGGTTTGGGATTAAAGCTTAACGGTGTTTTAGTCTGAG
CCTCAAGAGACTTTATTCATACTCTTTAACCTCTATAAACCTCTGCTTCTGCACTCCAA
AATGGGACTAACTATACTACCTTGCCTACTTCTAAGTTTGTGTGAGTATCAAAATGACAA
AATGTGTATGAAAATGACAAGCCAAATGTGTTCATCCTTGCAATTACCATAGTTGGTTAGAA
AGATAATTAAGCAACGGTTAAACTTTTAAATCCTGTTTACTTGTTTTTATTATGATAGAG
TTCCTTCTCTATTATTATAGAGTTATATATATGTATCATCCGTATCCAAACTAATATT
AATGTTATGGCAACTGTGCTTTTGTAGTCTTTTCCAAAAGATTTCACATGATTTTCATAA
AGCATATCAATTTTTTCACTCACATTTGTGTGACATCAGGTTAAAAGTTACTAACTTAA
ATTTCTTAGAAATATTGACTAGAAAGAAATATATAACAATATAAAATGTAGAGACGAGTATG
AAATATATTTAAAACTCATTACGATACTTTAGTTTCATTTTGCTATTTTGTACAAAGGAAAGCT
CAAGTAATTTTCAGTACCTATGGCAAGATTTATACATTTCTGTAAATTAGATATCTTTTAAAA
CCTACGGTGATAAATTTTCTTAAAAATTGGTCTTAAAGTAACATTTGTTAACACATTTGGTAA

TAAGAGTATTCATTGCCCTTGTGTTTCTTTTCCCTCATTTTATATACAATTTTGTTAATG
ACTGTAGATTTAGTGGTATATGTGGTAGCTCATGTAAATGCTCTAGTAGCTCAGTTTAAT
TATAGATTTTAAAGGATAAATTGCCAAGCTTATTTGTCTGCAACAACATCATTTGGGA
ATACTTTAAAAAGTTTTTCAGGGAATCAGAGAAAAACCAAACTAAAGATAATGGCCTCAGAA
CCCAAAATTATTTTACTCTGTCTATGAATTTTGCTAACAAAAATCAAAATAATGAGTACTAA
CTGACAAGGACCATTTAATAATCTACCCAGCTAAGGTTAGCAGTCAATAATTTTACATGAT
ATTTCACTGATAACCTTATTTATCATATTTTACATGATAACCTTAAGATTTCTGTGGGTCTTA
AGTTATGCATATTCTCTTTTATTACTTTAATAAATAATTAACCAACAACATGATGTAATTT
TAGCTTATTACTAAGTATTATAAGAGCACTTAGAAATATAACTTTTACAAATTACCAAAAA
TTTTTTAATGTCTTATCAACATGAATACTTAGTCATTTATTTAGAAAGTATAAGTTTAT
TAGTGGATTTACTGTGTCCTGATTTAGGTCAGTTACCTTAAAAAATACGTTTCATTTGTCTG
TTTTTATTAATGTACATGCCAAACTAATGACATTTAAAAATCATAGGAAAAAGAGATTTG
GCATTTGATTTATATTTCAATGGCAAAATGCACAGGTTTGGGTAATCTCAGGTTCTTAAAT
TCTTTCCACATTAATAAGTGAGCTAATTCIAACTTTTTTACTGTACTTCTTCGTGTTACT
GCATTAGGTTTTTGAATCTGCTTGGATGATGACTCCCGAAGAGCAGTTTAGTCTTTGAGAAAA
CCAGATCTCCATTGCTTCCATGAAGATCTTATTAAATATTTAGATAATGGAGAGTAACGAA
AAGAAATTTGGACATTGAATAATGCTTTGTCTTCTAGCTATCATTTCTCCCCAACTTAATCT
GTGACCCTGAGAAAAGTTACTTCATTTTCTCTTGAGTCTCTCAACAAATTTTCTTACACAGT
ATATGCCATCTTAATTATCTTAAAGTATGTACTTGATCATCTCTAGAGGTTCCCTTTCTAGT
CTTAGGTTGTAATAGTCTTTTGTGCATACCTGGAAAAACATTTTTTAAAAACCATCAAAAATGTA
TTCTTTTTTCTACACTGGAATTGACATATGGCACTATAAAGATCTTTGTGATGGTGAAG
TCATACATTTTGAACATGAGAAATATTAATGACAGCTTGTGTAACCAAGTATTTCTTATCCTA

TTCTAAGAAATAATTTAGTTATTAAGCTGGTTTTTTTTTGGTATGTCTTGAGTTTATTG
AACGATTGAAGTATTTGGAAATTGCGTCATGTAATAGTTCACTGC

SEQ.ID.NO. 388 HG1014923N0 PLT00014330C18.contig
CGCTGGAGACTCCGGGCCCTACTGTCTGTCTTCACAAAGCCAGGCTGCTGCAGGAACA
GTTCCCGCGTCTCTAAAGCCGGGGCGGAGGCTAGTGGTGGAAC TTGAGGTTGGGTC
TGGAGGAAGAGATGCCTGGGCAGCGTGTCCAGCCCCCTGCCCTCGGGGTGAAAGGCAGAG
AGCAAAGAGGAGGCCCTCTTATTCCGGACCCCCAGGCCAGAGCCCCAGGAGTGGGAAAGC
AGAGGGCAAAGAGGAGGCCCTCTCTCATCCCTGACCCCCAGGCCAGAGCCTAAGGTTC
AAGTGCCTCAGGCCCAGTCCCCCTTGACGCTTCCTGTTTGGGAGTTGGATTTTATACCACG
GATTTTATAGCATTTTATATAATTGAGGATTGTCAGAAATTGGAAAGGGCCCTGGAGATC
ACTAGTACTACCTCCTCACCTTCACAGATAGTAAATTGGAGCCCAGGAACCGAGATGA
TCACACAGCCAGGTAGGGACAGTACAAGAACTAGAAATGAAAATCCTAGCTCCTGACTCAG
TGTTCTTTGCACGATGCCACACTGCTTCCCAATTCGCTGGTCAGTTATTAATAATTGTCA
GGCCTGCCGAAGGATTGACATGCTGGGCAATGCAGGCCACCACGGTGGATCTGTGACTCC
TGAAGGTGGCAAAGCCCTGCACACTGCCTTGGGATGAGCACACTGAGCTCTGCAGGGCCT
CAGCTGGAAGGGCCATCCAATCGTCCTTCCCTCCCCAGCAGAACTCCAGAGGTGGGC
GGCTGCCGCAGGGCCCAAGGCTGCCTCGGCAGGGGGCTTTGTGCTTTTGTGTGGAGGC
CACCAGGATTTCAGGAAGATCCTGAATGGGTTGTCAATTGACTGAAAAGTAAATCCCCAG
TGAGTTTGAACCTCCACGGTTTTTTGAGACAGTATTAAGCTTGAGAACGGAGATTATGG
ATGGCACGAGGTGGCACTGCCATTCTTTTACTTTCCTTCATGCATTTTTCGGATCCA
GAGTATTAAACACCACTTTTTTTTACTCTGCCCCATTCAAAAACAGTAGGCCAAAATTGGGGGAA

CCAATTTATAATCACTCAAGCCCACTGCCAAAGTTCCGGATAAGCCAGGGCAGAGGGGTTA
 GATTCCCCCCTCCGTAGGCAGAGAGAGGCCTCAGGAGAGTGACATCTTTACTGAGAC
 AGCCCCAAACACACCAGTGGCTCTCAAAGTGTGTCCCTGAGACCAGCATCACCTTAG
 GAATCTGCTAGAAATAAAATTTCTTGGGCCATACCCCAGGGTTACTGAATCAGAAACTTG
 GTTGGGCCCTAGCAGTGTTTTCATGAGCCCTCCATTGATCTTGATGCAGCTGAAGTGCCG
 GGCACCAGTGAAAGGGAAAGGCTGCGGTAGTGATTCATTGATTCCTGCTTTGATGGAAG
 TTATGGCCTCACAGGTTAGACGTTGCAAGTTGAGTTGGCCCTCTTGCCCTAGGTGTGCCTC
 CTTGCCCTTGGATTCTGACATGAACAGGTGAAGTCTCTTCAATGATTTTGTGTGTGT
 TGTTAATGGTATGCTTGCTATTATGGCCAAATATGTAAATGTTACTCTGAATACATAATTA
 TATACCATGTTACTTAGTGACATTATTCAACATTTGTATGCGTATTGTGAAATTGTTT
 GCATCAGTCATTTAAAAATATTTCAGCCTAAGTTTCTAAAGAGGTATGTTGAATACATT
 TTTAAAAGCCAGCTTAAGGGTCAACAGCAGCCTTAGTTTAGGAGAAAGAGCTAATCACAGGT
 GACCACTCAAGTTCCTTCAATGAAGACCTGATGATGTTAGTTCCCATTTTCCACGGT
 GATTGACTCAATTAAAGTGATTTGTGGCTATATGTAAATTTACAGGCTGCTTTGCCGTTGCA
 GGTGGGAAATCTCCAGGCTGTTGAGGCTGAACCTGGGTAAAGACCCCTGACCCC
 TCCTCATTTCTTAAGGTCACCACCTCTTAGAGGGTCTCTTATTTTCTTTCAGAGATCAA
 ATGCACAGTTATGCTGTACAACTAAATTCTGCAGTCTTGGGAAATCCTTACGAGTACCAAC
 TATTAGTCACTTTTTAAAGTCCCCAAAGTGGAGTGAGGATAGGGGGAGTGGAAGCAGTGT
 TTGGTTCTTACTTACCCAAACTGTGTATAAAGAATA

SEQ.ID.NO. 389 HG1014925N0 PLT00014330D03.contig

GCCTTGCCACCAACGGGCTACGAGCGGTTAAGAGACTACAGCTCCAGGATGTACGGTGA

TGGCA CCCGTTTGGACTCCGCCCTCTGCGGCTCTAGAGCTGTTCTCCCGTGC GGAGGGG
 ATCCTGGCTGTTCCCAAATCTCGCCTGCCCTGCGAGGAGGCCAGCGTGGCCCGGCAGCTGA
 TCTCAACCTGTGATTGTGACATCACCCCTACCTGAGGAAGCCCTCTTCCAGCCTGCTCAAG
 GTTGCCGTAGTTAGGTTTACCCCTCCCTTCCCTTCCCTCACCCCGTCCCTTCCCTCATCCCT
 GCCCTTCCCTCTGTCCCA CGCGGCTTTCAGACATGCATAGTGCAGCCCCCTCAGTAAAGGT
 GCAGCTAATGGCCCGTGGCCCTAGCAGTTGT CAGGCCCTTTTGGGAGCCCATGATGGGAGC
 TCCCTGGTAAGTGAAACCGTCAATTAACTGCCTAAATGACAATAACCCCGTCTCTGGTAT
 CTTTGCACTTGCCCTCTTCAAAGCCACCTTTTCCTTTCTCTCTCTCTGGACTGTACAGAT
 CCTCCTCTGTACCCCAAGAACACCCCATTTGGATTTC CGGGCTGGCTGGCCTACACA
 CCTGGGTCA GGCCTCTCGCAGGGATGCGCCTGCCACTGTAAATAAGAGGAGAAACGTCA
 CACGGGAAAGGCCTGACTCCTTCGTACAATCAGAATACTCAACTGCACAGAGAAGGACC
 CCTTTAAGCCACTTTGGGAGCCACATCCACC ACTCTGTGACTCCCAACACAGGCTGGTTCC
 CAGGTATCAGGTGTCCTAGTGTTAACACGGGCCAAAGGACA AAACAGGACCCCTAGTCC
 TTCCCAAAACATTAGGAAAGTTTGTGTA CTGAGGTAAATGTACAAATGTCTTTTTGGCCCTGA
 ATAGGGCTCCCTTAAAAACTCTTTTAAATAATCAATTTTAAATATACTCTGGAACCACTT
 TTAGTAATTTGCTGAAGGAACATAAAAGTAATTACCTTAATTCATGTTTTCTTTTAGTT
 TTTCAGATTAACTAAGTAATTCATAATTGTACATTTCTGAAATTGAAGTACGCACAGCTTC
 AATTCTGTGCAGGGAGGATGCAGCAGGAGAGAATA TCCTGGGCACATCTTTCTGGAGAAT
 CACTTTTACTACAAGATTGTAGAAACAAATTTATTTCCAAGTAAGAAATTAAAACTC
 TAAACAAGGCTAGAAAGTCACTGCGCTTTAATAACCCTTAGTCATCTGTGCGCTGATATG
 TGAGACTTTCTCCAAGATAAATCTCTCAAGGATAACCTATTTTCCATCATTAATAATTAGT
 AAAAATTAGAAATTTTAGAAAGTCAATTGGGAAGCCCTGTCTGCCAGTTTCTTAAATCAC

AATGTGGCCTTTGTACGATTTTACTTGGATTTTCTTTTGGAACTCTGTTGACTGCTTGGGA
 TATAAAATGTGGGGTTTGTACAAATTTTGTGCAAGTTTTCAGCTATTGTTACTTTTGCACC
 CCATTTAATTTTCTGTCCCTTCACCTCTTTCTAGGACACAATCTGCCCCACAGATTGTAAG
 ACTCTGTTCAATTTTCTTCAAACTTTTCTTTTACTCTTTTCTTTTTCAGATTGAATAATTTCAATT
 GTTCTGTCTTTTCTAATCTATGAATCAACCTAAAAGTACTATTTAAAATTTTCATTTATCT
 TTCTAATTTGGTTCCTTTTAAATAATGTGCAATTTCTCTGCTGAGATTCACACATGTATTCAATT
 CATTATGAGAAATTTTCTTTCACCCCATGACTATAGTTTTAATAGCTGCTCTCAAAATA
 CTTGACTGCTGATAACAACAATCTTGGACATTTTGGGGATAGCTTTTAATGCCTATGTTTT
 ATCTTGTGTATGGATAACATTTTGTGTTTCTTCTCATGTCTCTTAAATTTTAAAAATTGT
 ATTTTAAAACCTGTAAATAATACTTATAGAGACTCTGGATCCTGTTGTATTCCTTTGAAG
 AGTGTGTTAGTTTTGTGAAGAGGGAGTTAATTTGGCTGGATTCAAATTCAAATGCCTGTC
 TCCCTTTCGGTGGCACAGCTAAAATTATCATTCGCTTCTTATCCCCACATATATCATAT
 GTATGATATATAGATGTGTGTTTCTATATAATATATGACATTGTATCCAGATTTTACCAT
 TATTTGTAAGAGTAGTGTTCACAACAAGCTACTCCACTATTACTGAAAGCCAAAACCTCAAT
 TTTATATTCTTTTGGATTTTACATAAATGACATTATATAGTATGTATACTTTTACATAT
 ACTTTCTTTTGTACATCATATTTTGTAAATATTCAATCAATCCTGCTGTAGATATGTAAGTAC
 TGAGATTACAGGCATGAGCCAAACGTGCTCAGCCTAAAATATTTTAGGAGGTTAAGGTGG
 AGGATCACTTGTGGCTGGGAATTTAGACCAGCCTGGGCAATAGAGTGAGACCCCTCTCTC
 TACAAAAAATTAAAAATTAGCCAGGCACGGTGGCATGTGTCTGCTATTCCAGCTACTCA
 GAAGACTGAGGCAGAAAGATCAGTTGAGCCCCATAAGTTCAAGGCTACAGGAAGCAATGAT
 TGCACCTCCTGCACCTCCCATCTGGGTAAACAGAGCAAGACTCTGTCTCTTAAAAAATAATAT
 AAAAAATTTAAAAATAATTTTATACATTATGTTAAAAATACACATAAAATTAACATATTTTAAC

CATTTTAAAGGTTCCATTGAGATGAATTAAGTACACTCATTAATTTTGCTACCATCATTT
 CCATATATAAAAGCATATTCCATTTTGTGAAACTGAAACTGTACCCATTAAACAACCTCC
 TTATTCTCCCTGTAGCCCTGGGAAACACTCTCCTACTTGGTGTTTTCTATGACTTTAAC
 TACTCTTAAGTACCAACATATGAGAAAGAAATCATATAGTAAAGAAATCATGAGGAAGAATGA
 TAAAAAATGTATCACATGCTTATTAAATTTTCAATAAAAAACACCAACAGAGATCAGTAGTA
 CCTGTTGATTATAAA

SEQ.ID.NO. 390 HG1014927N0 PLT00014330D05.contig

GTATTCCAGGTAGAAAGAAAGCAAATGTGTGGCCTTGAGTGGTCACCTCTCTTTGGCTTAT
 CGTATATATCTATGAGCTAATGAGAAATGAAAATTCCTTTTTTTTCAAAATIGTTAGGTTAG
 CTTAAAGAGATATTGACTCATTAATAACAAAACACAAGAGACATTAATATAAAAATTGTA
 TACATTTTTTAAAGCCTATATTTTCTCCCTTGAAATTAATAAAATAAAGTGCAGCATGTTGAA
 ATGTTGAACCCCTGGCCATATGACAAAATTGGGAAGTCTCTCTTTTGCTTACACCTTGGGGCA
 GAGGCAGGCATCAGACAAAAGTCCTTTTTATTTTGTATCCTTAAGAGTATTTCAAGGTCAA
 GTTAAATGAAAAACTAGTCATTGGCTCTGTTAGATTTTCAAAATATATCCTCAATAAAATG
 TACACTATTAATCATTTTCATATTGGTGATTTTAATCAACTTTTAAACCATTACTTTTTTA
 GAAAAGTTTTTCACAAATTTCTGGTTTCTAGGGTAATTGAAATTCAGTATATAATTATAG
 TAGTATAAATATGATACTTAAAGTATTATTAAGGTCAATTAATAATTTTAAAAAAACT
 GGTAGGGCAAGAAGACTTTACTACAAGTTAATGAAATGGGGGTAATTTAGTTGAAAAGAA
 ACAAAATGATTACAAAATTAAGATTACTATTCTGTGGCACTCTTTCTTTAGTTGACAGTG
 AATAAGGTAGAGTACTATATTTTTATTTCATATTATGATTGGCCCCCTTTCTAGAAAAAGT
 GAAAATTTTATGTAACTCTGATACATGGGAATATATTTTGAAATTAAGAAAGTATATGTACA

TATATTAGTCATTATTATTACTGCAACTATTCCCTAGTGAAAAAATATACCTCATTCAAAGA
 GAAAGATAAAATTTCAACAACACTGTTATATATGCACACTTTTGGCCCATAAACTAAGCAAG
 CCTAATAAATTAATAATTTATCTAAGTCACAACCATCCAAATTAGGGACACTTTCITTGCAAT
 CAGATTACAAAAGGATATTTTTATCTCTTTTCAATTTAGCAGTCAGGTTAGGGCTCCCTAAT
 AGAAGGCTGCGTAGCTCATTAGTCCTACCACTTTGGATGGCCTTTACTCTTAACITTCATT
 CTCGGTAACTCAAGGGGAGTCAGTGTGTAACTGTGGTGCAGAGCAATGAGGATGCT
 TGAGTCTTCATTCTGTGAAGGGATCACCTGTGGAAAGAAATGATCAGCTTTCAGAAAAATCTCA
 TCTGGTTAAGTCATATTTCTCTGCTTTAAAATGTTACAGGATTTCTCATTTGCCCTTAGCA
 CAGTTTCTCAAACITTTATTGAACATCAGAAATTACTTGAAAGGGTTTTCGAACTCATGAATT
 ATTCGGCCTTGCCCCAGAGTTTCTAATTCAGTAGGTCTTGGGTGGGAACATACTTTGAG
 AATTCTGACCTATAGGATAAACACTAGAGTTCTTATTATGTCTTACAAGGCCCATCATGA
 TCTGAAGGCAGCGGTGTTGCTGGTGGAGGATGTGAGCTTTGGAACTGGAGCTCTTTGAT
 TCAAAGCTCTGCCGCTTACTCAGCTGGATAAGACCATGATCAAGTTATTTTCATCTGAAAAA
 TTGAGTTGACTGCAGGTAACTGAAACTGCCGAAAGTGAAACTGCAGATAAGAGGGACCAC
 TGTAATAAAGTGACATGAATGATACTGTCAAAGGACCATAGCCTAACCCAGCCAAAGCAAAC
 GGAAAGTACCTTTAGACTCACGCTTTGCCAAAGCTCAGTTCCAGAAATCACATGATGAAAAAGA
 ATGATGGAAAGAAAAATGTTTGATATTTAAAAATTAAGGAAATTACTCCAACTCAGAGTTT
 CATCCTAACTGAGAGACCAACATCAAAAGAATCTTCTCAAGTTTCTGAGTGACGGCAAG
 AAACTAGACTACTGTTACCAAGGAACCTTTTAAAGCAAACTGCTTTTGCTGATATGAGGA
 GAAAGTTAATTAATAAGAAAAATACCTTAATGTGTTTGAAGCTGTTAAAGGTGTGTAAAC
 AATTACTCTGAATTTATTATAGCACCAGTTAGAGACCCAGGAAAACTATAAAATTGCACA
 TCAGGGCAGAGCTGAGAGACGACATCTGCTTTGGGATTGAAAAAATAAAAAA

AAAACAGAAAGGGAACAAGAGAAATAGCAAGTCATTTAACATTTGGAAATTCAGGACTCCA
 GACTCATCTATAGGACCAAACTGCTCTTGAAATATACTAGATTTTCCTTTAGCTCCAGAGG
 TTTTGCCTTCATCAAAATGCCTGTAAACACTCTTATTGTCAAAGGAATAAGTGATGGTGAA
 CACTGTGCCAGATAGGAGGTACCAACACACACAGGTTTCAGCAAAGCAGAAAGCAAGAGGTAT
 ATTTTGTGTTAGATCTTAGTTTCATCCCTATGTAAACAGACATTTTCAGGGAAAGGAGGAGAA
 AATAACAAAGCTACCTATGAGCAAAGGTAATGGGATTCCCTATCCTTGTGTTTCCCAGTTT
 CGGTTCATGGTATTGCTTATCCATTATTAGTTATCCACTGCTGGGTAAACAAATTAATCTCCA
 AACTTAGCAGCTTCAAAACAATAAAACATGTATTATGTCACATAGTTTGTTCAGAACT
 TTGGGAGTGGCTTAGCTGAATGGTTGTAGCTCCAGGTGCTTTTGTGGTTGCAGTCAAGA
 AGTCAGCCCATTCACACAGCTGCAGACAAAAGGACTCGGTTCCCTTGCAGGCTGAGGCCCTG
 ATTAGGCAGGGGCTGAGTGTCTCGCCACATGGACCTCTCCATAGGGCTGCTTTTTCGTG
 GCATGGCAGCTGACTCCCTCAGAACAGATGATCAAAAGCAGCAAAAGAGAAAGCCTCACA
 TGACTTTTGCCCTAGCCTTGGGAATCACACTCTTCCATTCTACAAAATCCTATTGGTTA
 CATAGCAGCCCTATCCAGTGTGGGAGGGAACCTGTATACAGGCATGAATACCAGGAGTGGA
 AAACCACTAGAGACTGGCTACCATGGTCACTTACACACCAACATGAAACTTTCAGCCTTT
 CCACCTCCTACCACTTATACTATATTTAAATGTTTTTCTGTTATGGACTTTAAAAAAGAG
 GTCATCATAAAGCAATAATCCATGAAATCATGGATATCATGTGATCTTTATATTATTT
 TTATTTTATATATTAAATTTCCAAAT

SEQ.ID.NO: 391 HG1014929N0 PLT00014330D07.contig

TGGACCTGGTCGCGCTGTGCCGGGTGTGACACTACGGTAGGTAGGTAGGATGTGCT
 TCTGTCTGCTACTAGTTGGGGCCGAGGGAAGAAGATGACGATTCCAATTTCAGTAG

CCACTCTCTAGGCTTATGA ACTGAAAGACCAAGGAGCACAAATTTCCGAGTTTCAAAGATA
CTTTCAAACTGAAAGGA ACTGTTCCGCCAGCTGCTCGCGAAAGTGCCCCATGTGACCCCTAG
TGAACGGTCTAAACTTCCGCTAGATGCTTTGGGAATTCGTAAAGCGACTGCGTGCGCGGC
AGCCAAGCCACAGCTTTACAGGGTTCCGCTTCTGTGTGGAAACGCAGCTTTTCCCGAGAA
ATTGGGTTAGGGATTAAAGAGAAAGAACTGCATCCCTTTTACTGCCCTTTTTCACCGTCTG
TCCTGGGCTTAGGGCCCTCCTCAGAGGCCAGAGGGGTAGGCGAGTCCGGGCGGTAGGGG
GCTGCGTCCGGAAGTCCAGGCTGCGTTCTGTCCACCCCTGCCCCACCCCTTAGGGGGCAGGAC
AGGAATAAGGTCACCTCACCAAAAGAGGCCCATCAACATGGAAGAATAITTCAGTGAATCT
TGACTCACCAAGCTCAGTGACTTTTTCATTGAAAAGGGTGAGAAATTATTAAGGAGATAATTG
CACAAATTAATGTTTCATGTCAAGTTTGAAAAAAAATAATTCITTAAGCAGCTAGGAAAGAT
ATTTGTAA AATGACGCCCTGAAAAACTTTTCATGTCTGCCGGACGTTATATGAGAAAAATACTT
GAGTGTGTTAGTTTCTATTGCTGCATAACGAATTACCACAAATCAGTGGCTTCAA AAC
ACCTATTCTGTTTTCTCAGTCCCTGTAGGTCAGAAAGTCTGGACGTGGCATAAGTTGACGGTTT
TGCTCAGTCACACAAGGTTGAAATCAAAGGGTCTGCAAGTGGGATCTCATCTGAGGTTTG
GGGTCCCTTCTAAGCCCTCTGGCAGAAATTCGTTGTGGTTGTTGGACTGAGGTCCCCTTTT
GTCTCGCCAGCTGTTGGCCAGAGGCCAGCTGTGTGTCTAGCTAGACCTTTTACAGGCCTTC
TAACACTTCCCATCTTCGTGGCTTTCGGGAGCCCCCAGCCCCCTTTTAAAGGCTCACCTGA
TTAAGTAAGACTCAGTCAGATCTCCCTTTTGATGAACCTAGTCCCAAGAGTGATA TCTGC
TCATAATTTGTAGGTTCTACCCAAACTCCAGGGGAGGGGATTATGTAGATGCGCTCTCTAG
GGGCCAGGAATGTGGGGA ACTAGCTTAGAAATTCCTGCCCTGTACAGCGTTGTATATATCATG
TTTCGATTAGTTGGAAAGAGTAAGGCATGGGTGTGATAGGTAAACAGGTTGACTTGACCCAG
TAAAAATTTTTTCCTTAAATCCCTTAAATCTCTGTGTGTATAAATCACAAATCTTTTGATA

TGAAGAGCAATGTCAGGGAAGTGAGGAGTAAAGAAAGCAAGGGAGTTGGGATATCCTTCAC
 CCCTTCCACTGGAGCTTGGGTAGAGCTAGAAATTTTGGCGTCACGCTGTAGACCATGGT
 GCCCATGTCTTTCCACATAAAAACTCCCTCTCTTCTAAGGAAATTTTATGTTAAAACTCA
 GCACTTTAGGAAAGGTTTTCTTGAGTTTCATGATGATGTGCTTTTTAAAGTTACCTGAAGC
 TACACATAAATTGTTAGTATAACCTGGAAATTCTAGAAATTAAAAGGCTGGCACAAATTACG
 ATTTTCTATGTACAGGGAAGTACTAAATATCTGTACTCTATGAGGAGACATAAGACAT
 GAGGTTTAGAAGGAATGGTCTGTGCTTTTTGAGGAGTTTGTAGTCTACTTTGGTTTTAAAT
 ATTTATACTTTTTTTTAAAGGAAAATTGGAAGCTCATGAGTTTGGCCTTATCTGGTTTC
 TGCCTGCCCTGGGAAGCCCCCTATCAGATTACITTTCCCTCATCTGTCTGCTTCTTTCAGTAA
 CTGCGCCCTCCAAATGTTCTTTTCCACCCTTAGTGCTTTCCATAAATGTATATTCCTTCTCTA
 CCTAGAAAATTTCAITTTATCTGCTCAGATTCCTAATTATATTCACACTCTTCTTAACTT
 AAATGTTACTTTGCTTAAAGACATTTTCCCTTGGCCCTTCAGACTGTATAAGGACCTCTCT
 GAAGTATTTTCAATAGCATCTCATAGTGCTATTTTCAGAGCATTTATTTTTATTGTAAAAAAT
 TGTGTAATGGTTTGTGTTATACCTGATCTGTGTTTTCCCTGTATCAAAAAATAAAAATTAGGATG
 AATCTAAAAGTTTATTGTCAATAGAAAAGTACAGATCAGAAATCCAGGAGACTTCAAATCA
 AGTGGGAAGAGTGGTAAGAAAGCTTGCACTCACCAGTAACAGCACAGCTTATAAGGATAAA
 GGAGGAAGCATTTTGGAGCTTTTCCATTATTGTCCTTATACATTAAGATTTTGTGGGG
 GTGGCAACAGCTGTTTAAGTTGATTTGTCTACAGTTGATTGGCTTAATTTACGGAAATC
 ATACTAACAAAAATGTAAAGCTTATGTTTTATATTTATGATTAGAGCATTTTCAGGGAAT
 CAGGATGACTTAAGTTTTGGTTTTATGTGGTTATGGGCAGTTGGTCTTGATGTATATCTAA
 ACTGTGACTTCCATTTTATTTTATTTTAAAACTGTAAAGATCCATGATGGCATT
 TAGTGGGTTCTCAACAAATATTTGTTGAATGAATGAATATTTTGCTTAGGAAAGCTTTATG

TAAGAAATCCAAAACATTTTCTAATTGAAGAGTAGGTTGGATGGATGAATGAGGTCACCTTT
TAGAGAACTGTGAAATACCTATCTGTGTTGATTGAATTGAGGTACCAGGGTTTCTAGGT
TCCTAGGATTGACATAATTAAAGGGGCATTTTTTTTTTCAAAATGTTAATTCAATAATGGT
TGTAAGAAATGGATTGTGTACAGGACAGTCTAGGTTAGTGGCCTCAAACCTTGATCATGTAGA
AGAAATTACCTGTGTGTTTGTAAATTCCCTGGCTTTTATTCCACCTTCCCCAGAACTGGTT
CAGGCCACACTTAAGCACCTGAAGTATGCTGATCCAAATTGTAACATGCTGCAAGTGTA
AATACCAGATTTCAAAGACATCTAAACTAATGGATATAAAGTATTATGTTAATATGTTTT
CTAATTGATTACATCTTGAAATAATATTTTACATTCAATTGAATTA

SEQ.ID.NO. 392 HG1014930N0 PLT00014330D10.contig

GGACCTCGAGCGGCATGCACAGGTCCAGCAGGCCGCATAGGAACTCCACGCGCTGAGCGG
ACGGCAGCTCCGAGAACCCAGCGGATATGATCCAAAGGTGCTGAAAACTTCCTAACAAAAC
CTTCAGCCGAAATCTGGCACGCCCTATGTTTCCAGCTACTGGGTAGACTGAGGCAGGAGGAT
CGCTTGAGGCCAAGAGTTCAAGGCCAACCTGGGCAACATGAGTCTTTGCAGATAATTAAAG
GTAAAGAGTCAAGATAAAATCACACTGAATTAATGGTGAGCCCTACATGCAATGAGACAGTC
CTTACAAGATACAGAAAAGAACACATAGAGACTTCGGGGGAAAGGAGACTATGTGCAAGG
ATGGCAGAGATTGGAGCCGAGGGGTGCTTGGCCTACCAGAGGCTGCAAGAGACAGGGAG
GACCTCCCCTAGAGCCTCCAGGGGGAACAGCCCTGCCACACCTTGACTTTGGACTCCT
GGCCTCCAGAACTGTGAGGACCTAGGACCCTACGAGAGGAGGATCTGGAGGCCAGGA
CAGGAGCAGGACAACAAACAGAGGCCCTCAGAAAGAACTCAGAGTGGGACAGGGGTTTC
CTTAAGTGACGGGGGCTGGAAAGTGCACAGAGAAAGCCTGGTGCAGCAGGGAATTTCTCTC
CCTGGGGTGCTGCCAAGGTGTCTGTGTTGCTTCCAGCTTTTCCAGCCATGACTGCAGACAA

AGTCCACCTAGAGTCATAAATGGGCAATCAGCCCCGAACAACGTCACGCCCGGTTTCTG
CACGGCGAGGCGTTTTTGTCCGTGGCAGACCTGCTTACCCGGGTGTCTTTGGTGTGGGC
TAGATGCAGCATTTCCCCAGTGGGAATGGCCTGGGTCTTATTTAAAGTGGAGGTCCTA
GGCCTGGCCCCACGTATGGAGTCAGGATCCCTAGGGTAGGACCTTGGAATCCGCAATTTT
TTTTGAGACAGGGTCTCACTCTGTCACCCAGGCAAGTGGCACGATCTCAGCTCTCTGAAAC
CTCCACCTCCTGTTCAAGCGATTCTCTGCCCTCGGCCTCCTGAGTAGCTGTATTACAG
GCACGCACCAACCATGCCTGGCTAAATTTTCTATGTTTTTAGTAGAGATGGGGTTTCATCAT
GTTGGCCAGGCTGTCTCCAACCTCTGACCTCAAGCCATACACTGGCCTCGGCCTCCCTA
AGTGCTGGGATTACAGGTACGAGCCACCGTGCCTGGCCAGGAATCTGCATTTTTTGCTAGT
AACCCTGTCTCATCTAGTGCACACCAAAGTTTGAAACCACTGTTGGGTGTGGAGTTCTT
GCTTCCAGACTTCCAGGGCCTATGGATTGAACAACAGCAGCCACAGTAGTAATTAGAATT
TTCAGTCTGAACCACAAGCCTGCCATGCAACAAGGTTTCCTCCGTTTACAGATGAGG
GACCTGAGGCTTCATAAAGGCTGTTATTGACCTGAGGTCCGACAGCTGATGGCTCCTGGGT
TCCTAGCGTTAAGTTTGCAGCCCTTATTTCTGTTGTATAATTTGTTTCTCATCTTTCTGTG
TTTATTGCACTGCAATTGTGATATCCTCTAAAAGCCAAGGATGAACCTCACTGCTGACCTT
GGTTTTATTTCGTTTGTTATTTTTCACCTTTTCTTGTAATGTGCAAAACATACTCAAAAG
TCAAGAGAAATCATAATGGTGATTTCTCATAGTCCCCCTTTGGCCAGCTTCAGCAAAGATGAC
ATTTTGCCACTCTTATTCATCCATTCTGTCAAGCCACCCCTGACATACACACATGTTT
TTTCTGGAGTATTTCAAGC

SEQ.ID.NO. 393 HG1014932N0 PLT00014330D12.contig

ATACTGACCCCTTTTCTTCCAATAAAGTTATTGCAGTAATCACATCATTCAGAGAAATA

TTCTTAAAGCTTTTTTCAATAATTTACTTCTTTAATGAAAACTTAATCCAGATTAAAGC
 CACATACATAGAGCTCTTGACTTTTATTATTTTGAATACCTGAGATTTCCAAAATCTCTTT
 AACTGTCTTCACACAATTTCTAGTCTGTCAATTTAATAATTACAACACTGTTCAATTTCTTC
 TGCTTTTGTTCCTCTTCATGTGCTGCCCTAGCGTTGCTAGAGTTTCAATCAGTTTCCCTGCC
 TTGAGAGGTTTTCTAGTACAAAATTAAACATTTTCCCTTCCTAGTTTATGCAAGTTCCCTTCCT
 GGTAAATTTTATCTTTTGTATCTGTTTCAATTCCTGCAATGTTAGTTAATTTGGAACCATTTCTA
 CTGTGCTCAAAATGAACAGTCCTATCCCTAGTTCTCTTATTCCTTTTTCATTTTCATATTTT
 ATACCCCTTTTCCCTCAGTCTGTCTTAAAGAATAAAGTGTTTTCATTTTCAAAAGCTTATTA
 TCTCTGTTTTTAAATCTTAGCCCTTCTTTTCCCTCTACTATTTTTCCTTCCTCTTAAATGCTCCA
 TGCTGGGTTCTCTGCCTTATCAGACTTAATCTTGAAAAGTACGTTACTTCCTTCCCCATTA
 TCTCCAAGGTTGTTATAGTCCATTTCTAGGTGTCTCCAATTTCTCAACTCTTTACTCAATTA
 CTTTCTGAAGTTTGACTATTATTTCTACTAGTGTTCAGCTTTCTTAAAGTGAAGTAAACA
 CATTTCTTACTTAACTAGGTAGATAAACATTAAGCTGGACGGATCTGGCCGTAAACATGCT
 ATTATTTCCAAGATAAAACTGAGTTCCAAATGGCTAAATCTTTGTCTTTTTCACCTTTTCAT
 CCATTCCCTGTCTTTTTAGCATTTTGTTTTTTTGTCTGCTTAACATTCACCTTCCTTTGTGTC
 TTCTGTAAATAAACCTTAATGGTTTTTCTCTCTCTCTGGCTAAATGTTTGTTCCTTTGT
 TACTGCTTTCTCTTTTATATACCACTTATATATAAATGGGTTGAAGCTTTTGCTCTTAG
 CTCCTTTTTCATTACACATCATTTTGGAGGAGATGGTCTCGTGTTCATTATGTAGATTATCA
 CAAATTTGATGACTCAGAAATACCACATTTTCTTAGTGACCTGTTTACCACCTTTCICA
 GTCGCCAGGTCTTAGCACCCCCAAAATTGTTATTTCCCAAGTGGAAGTTTGTCTTTGTACA
 ATGCTTGCTATATGGTTGGAACTAAGTAAATGCTTTTTTGAATCATGAACAACTCCTTAT
 CTGTCTTTTTTTTCTTAAGATAAATGGTTTTTAACTGTCCACCTATTCTCATGTGCACAAGC

CCTGTAATAATCTCTGACCACCTTCTTGTCTACACTTGTTAATATTAGTTGTCAATATTTGG
 TCCTAAGTTTGACTAATGAGGAAGATGACATTTTGTCTGCAGATGGTGAAATAAAAAT
 CACAGAGATTGTGATTTCCCTTATGGTTTGTGGGTAATGGTGGAGTTTAAACTTACAACG
 AAGTTTCTGGTGACATGTTCCCTAGTTTTCACAGAAAACATTCTTTTTTTTGTGAGACA
 GAAAGTCTCACTGTGTCGCCCCAGGCTGGAGTGCAGTGGTGTGATGTGTCGGGCTCACTGCA
 ACCTCTGCCCTCCTGGATTCAAGTGATTCTCATGCGTCAGCCTCCAAGTAGCTGGGATTAC
 AGCGCCCCACCACCGCCCCAGCTAATTTTGTATTTTATAGTATGAGATTTCACCA
 TGTTGACCAGGCTGGTCTTGAACTCCTGGCCTCAAGTGACCCACCTGTCTCAGCCTCCCA
 AAGTGCTGGGATTACAGGCATAAGCCACCACACCCAGCCAAAACATTCTTTATAATGAT
 ACAAGTAATATACCAGACAAGAATTAC

SEQ.ID.NO. 394 HG1014934N0 PLT00014330D13.contig
 CGTCTTCCCTTCAAAACTTGGAATGATTTCAAAATCATAGGCACCTTCACTTAACCCCTAG
 CTTCCATTTCAGCAAACACATCGGATCGATGCTACGCTAACCTATCGGGTTCTCTCTC
 CGCGCGTTCAGGTTAAATGAATACCTGACGAAAGGGCCACGTTTCAAGGCAGTGACATT
 TGATAGCTGAGAGGAAAAGTGGCTTTAATGAAAAGCAACCTTTGGAAATTCCTGCTTGTGA
 GAAATCCAAATTCAGCTTTTGTGCTGCCAGCAAGAAATGATCAGTAGCACCCAGTGTAT
 GATTTTGTGTCTCAAATGTGCGGCCTAAAGGAACCAAGAGCTACAAACCTAAAAGACTA
 GAAAGTGGGCAAGTGCTCACACTCTGGCAGCTCTGCTGTGCAAAATTAAGTGTGTGT
 GGTCAAGGAAATGAAAGCTACAATGTCTCTTTTGAACAAGTCGTTTATCAAGAGAAACATC
 TATCAGGGACTGCCTTGTACATTGATCCTCTAAAAGTTAGGTCGAGGTGGAAGCTGCT
 TTGATGTTTTTGTATCCTAAATTTAATACACAGAAAAAAGCAAGGATATTTTGACCAGT

ATAGGTAGTATTTCCAAAGAAATCTATAGAGACCATCTGTATTAATTACITTTGCATTTTCTAAG
 ATAGTTTAAGATATAAACACCTGAAAGATAAAATGGGTAATTTTCATTATAGGTTAATTA
 TTTTTCCTGAGCTTCTGGTTGTCAACCTTTTAGAAATTTAGTATCTTACTCAGGACATTCA
 TAGGGACAGATCCACAAAGTTTGAGAGTCTGCACCTGTTTTTCAGAATGTTCAATGTCAAGG
 TGAAGACAGCCATGCTAATACITCCCCTGTTTTTGTCTGTGTGAGTGTTTGGATGCCCTTGA
 ATGATTGCCCTAATTAAGTAATCACTTTTAGTGGCCCAAATCAACAGAAATTAAGATGT
 ATTTGAAACAATTGCAGGCTTTTAAGGTTTGAACTTTCTAGTTCITTAGAACCTTGGTTTA
 ATAGAAATAGAAAGGTGCTATTTTACCTTTGATCAATATATCTTGTGAGTGCAGATTTTCCAC
 ATACTACTCTTTATAAAGTACTGTGTTTCATTGTTGAAGAGCAAATAAATTTAATTACCAG
 GAATTTAAATTGATTTATAAAATTTTGTGATTGGAAGAACTATTCTGCAGATGTTGTGTG
 AAATATTGACATATGCTACTGATCGATAGAACTTGATTAAATCAGCATATCTGCTGTTTTT
 CTTGAGTCATTGATAAACTAGATAAACAAAATGGTTCAGAAAAAGCTATAAAATTTTAACT
 CTCATAAAAAAATAAATGT

SEQ.ID.NO. 395 HG1014935N0 PLT00014330Di5.contig
 TCCCCTGATTCAGAAACCACTGCCCTAGAGGATATGGCATAATGGTAACTTTATGTGATGGT
 TAGGGAATTTGAAAAAATAAAGAAAAAAGAAACAAAGGAATGGATAAGATTCAACATCT
 GTGAAGTGACACTCAGTAAATGAAAGGCGGATGAATTAGCATTACTACCAGCTGGTGTCC
 TACTTATGTTATAGAAATGTGACAGCAAATGTGATATGAACCTCCTTAGAAGCATTTTAT
 ACACATTGCTGGCTTTAATTGTGAATGCTACAGTAACCTCCTGACTGTTTTTCTCTTCTC
 TGCCTTTAGACAATTACCCCAAGGTTAAGATGCAATAAACATGTGCCCTATTTTCTATCCAA
 AATCTCAATCAGATTAAATGCTTCTCTCTAAAGACCTGAAAGGCTTATTTTGTTTATTTGTA

TTTCATGTAGGATTTGGGGGATGAAAGAAATTGGTGATTAAATCAACATATTATTTTCCCAA
TAAACAGACTATTCTGACAAAATGTTTAGTGGCAAAATAATAGCAACAGCAATAATGGCAAC
AATAAACTTAGAACAGAAAATGGATTTATTTTATTTATTTATTTTACCAGCTCTTTCT
CTTTCACTCGCTTCTGCAATGTTGAGATTTCTCACACATGCAACCTAGATCTTCTTTGTCTT
CTTTATATTCTTTCTCCAAAGGAAATCTCATTCACCTGTTGTGCCCTTTAAATAGCAAGCTTC
AGGTGACCCCCACATTTCTATCTCCAAATTGACATATCCAAATGCCTACTACTCATCTCC
CCTCTGATGTCTTATAGACATTCACATTTAATAGCATATTGTATTGAGTCTAAGGAGCCA
TTAATTGTAAGTCACATTTATGTCCACTAAGATTAAAAATACTGCCAACTGAACTATGA
CAAAATGCTGATATCCTTGATTCTAAGATACATCTTGATTTCAGAGACTTTAAAAAATAGG
CTTAAAAAATAATGTGATTTACTAGAATAGCACCCAAAAAGAAATAAAATACTTAGGAATTA
ACTTACCAAGAGAGTGCAAAACTTGTATTCTAAACTCTAAATAACATTGTTGAAAGAAA
TTAAAGACCTAAATAAATGGAAAGGTATCTTATGTTTCATGGATCAGAACACTTTGTAAAG
ATGGCAGTAATTCCTCAAGTTGATCCACAGATTTCGGTGCAACCCCTATCAAAATCCTGTCT
GGTATCTTTACATAAGCTTATTCTAAAAATTTATATGGAAATTCAGGAGCCCCAGAAATAGC
CAAAACAACCTTGAAAAAAGAAATAAAATTGGAAGATTCACATTTCCCAATTTAAAAAACTT
ATTACAAAGCTATAGTAATTAAACAGTGTGTTACTGGGAATAAGATTAGACATATAAATC
AATGAGGTAGGATTGAGAGTCTAAAAATAAACCTTCATATGTATAGTCAATTGATTTTCA
ACACAGGCACCAACAAATTCAATGGGGCAAGGAGTCTTCAGCAAAATATTACTGGGATAG
CTAGATATCCACATGTAAAAGAATAAAGTTGAATCCTATCCTCATACCATATACAAAAAT
TAACTCAAAATGGATAGAAGACCCTAAATGTAAGTGCTAAACATAAAAGTCTCTAGGAGAA
AACATAGGAGTAAATATGTATGACCACGGGTAAAAAAGCTTTTTTTTTTTTTTTTTTTT
TTTTTTTGAGACAGACCTTGCTTTGTGCGCCATGCTGGGTGCAGTGGCGCAATCTTGGCT

CACTGCAACCTCCACCTCCCGGTTCAAGAGATTCTCTGCCTCAGCCTCCCGAGTAGCT
GGGACTACAGGCACACGCCACCAATGCCCAGCTAATTTTGTATTTTAGTAGAGATGGG
TTTCACCATGTTGGCCAGGATGGTCTTGATCTGTGGACCTCGTGATTCACCTCAGGTGAT
CCACCTGCATCAGCCTCCCAAAAGTGCTGGGATTATAGGCATGAGCCACTGTGCCAGGCCA
AAAAAGGCTTTTTAAGGTATGACATCAAAAGCATTAGTGACAAAAAGAAAAACAGATA
AATTGAACTTCATCAAAATTAAGATACTTTGTGCTTCAAAAGATGCCATTAAAAAAA
AAGACCCTACAGAAATGGGAGAAATATGTTTGCAAAATTATATATCTGATAAATGACTTCT
ATCTAGAATATATAAGAATGCTTACAACTTCATACTAAGAAAGCAAAATACCCAATTAAA
AAATAGTGACGGATCTGATTAGCTACTTCTCCAAACAAGATGCCCAAATGGCTAATAAG
CACATGAAAAATGCTCACCAATTCTTAGTCATCAAAAGAAATGCAAAATCAAAATATGAGATA
GTATTATACTTCACACGCACTGGGATGGCTGTAATCGAAAAGACAGATGATAACAAAGTAT
TGGAGAGGAGGTGAAGAAATTTAGAAAACAGTCTGGTGGTTTCCCAAAGATAAACATAGAG
TCATCATATAACCTAACAAATTCCAATCCTAGATATATATCCAAGAGCAATGAGAACATAC
GTTACACAAAAAGTTGTACATTTTAAGCAGCATTATTTCATAATAGCCAAAAAGTAGAAA
CAACCCAAATATTCAGCAGGTGATGAACCTGAGAAACTGTGGTATGGCCATAAAATGAAAT
ATTATTTGGCCATAAAAGGAAATGAGGTTCTGTACATGCCTCAACTTGGATGAACCTTG
AAAACGTTATGCTAAGTGAAAGGAGGTCTTCCCTTGAGCCCCCATCTGAATGAGGAGTCC
TTGTACTTTTCATCCCTTTGAATTTATCATGCAATATTATGTAATTGTTTCATGAAAATGTTT
TACATATATATGTATATAATGTATATGTGTATATATAATGTATATGTGTATATATATA
AATTGTAGAAATTTCTCCTTTCCCATTTGATTCCTAAACAAATAACCTGTAAACACCTGTGC
TGTACTTACTATCTACCAGGCACTGTTATATATGCTTTTATATATTTTGTATTCATTTAACC
TTCTCAACAACCTCTTATGTAGCTGAGGTACAAAATGGCTAAGTTATCTAAGATTACACAG

CTAGTAAGTAGAAAAGCTAAGATTGGCTCCTGGCAGTCTTTTTCCTGAGGCTGAGCCGT
GGTCTGTGTATACTATAAGGGTGGTTCTGTATCTGTTCTTTTATCCCTGTGTAGCCTG
TATTAGCATAGCATCTGGCATGATGTAGGTACC

SEQ.ID.NO. 396 HG1014937N0 PLT00014330D17.contig
CGAAAGGGGATGTGCTGCAAGCGGATTAAGTTGGTAACGCCAGGGTTTCCCAGTCAC
GACGTTGTAAAACGACGGCCAGTGAAATTGCGCGCAATTAACCCCTCACTAAAGGGAACAAA
GATGTGTAACATAACGGTCCCTAAGGTAGCGAGTCGAGGTCGAGCTCTATTTAGGTGACA
CTATAGAACCAAGAGGGCGGGCCCTCCGGGGGAGGCTGGAGCGATCCAGACCCCTCGCAG
TGCAGCCCTCGAGCGGGCCCGACCGTCCGCTGCCCTCCCGCAGCCGCGGCCCTGGG
CCCAGCCTCCAGCCCGCGAGCGAAACCTGCCCTGTGCTTTGCTTCTTCTTCTTCCACAGTT
TTCCCAACCCTGAGAAAAGTGAGAGGTGCTTTCCCTGGTGTTCCTCTCAGGAAGGGCCGT
GTGACTGAGCCCGGAGTTGGGCGCCGTCTCGCCGGGCTGGAGGTACGCGGGCTTCGGG
ACACGCACCTAGGCGTCCCTCGGGGTCCCAGGCTCTCTGACCGGGCTCCCCCGGCACCGA
TCCGCGGCCGGAGCGTTTAGTCCCGGGACCCGTGAGCCCTGAAAGGGCCATTTCCCCC
TGAGGCGAGTCCCCCAGGCGCCCCGCAGCCTGGCCTCGGGACTGCCCCCACCGACAGTG
CGTGCGGGGGAGAGTGCACTTAAACTGCAATTTTCGAACATTAGTGCCCCGGACCCGCA
CCTTTTTCCTCCGTCTTTGATCTGTGCCGAGAGGCAAGGGCTGGAACGGCGCCCCCTCC
CCTGGTAGAGGGACGGAGACCGCAAACTCTTCCCCCTGGGAGGTCCGTTAAGTAGATC
CAAGCACCTGGAGGCGGAACCTGTGTGGCCGGCTACAGGGTGGTCTGTGTATGTGTTAC
CGCCTTATTGTTCCAAAGAACAGTCTTTTTCAGTGACAGGTAGAAAGACGCCCAATAAC
GAGAACTGACTGACTTGCCCAAGCTTAGAGATGGGAGTGAGTCAAGTTCCCTGGTGGCC

CCACGATCTCGCCCAACCATGTCAGCATGGGCTGCCCCGGTGAGCATAGCTGAGATGCACCA
 AGGCCACATAAGTCACTACTTTGGCTTGGCTGCCCCATCTGCATCTATCAAAGGAAGCC
 TTGGACACCCACAGAGAGGCCAGCATGAGGGAGTGTATGCTTTATACAAATTCACTCC
 AACTTCAGAAAACGTCCTCGCAGCTCATGCTCCCCCTGCCAGAGCAGCAGTGTGAGCAGCT
 GTGTAGGTTTGGAAGCACCCCACTCCTGGGCATTGATATGGTTTGGCTGTATCCCCCAC
 CCAAATCTCATCTTGAAATTGTAGTCCCGCAGTTCACAGGTGTGTGGAGGAACCTTGTG
 GGAGGTGATTGAAATTATGGGGGGGGTCTTCTCTGGGCTGTTCTCGTCATAGTGAATGAG
 TCTCACGAGATCTGAAGATTTTAAAAATGGGAGCGCTTGTTCACAAAGCTCTCTCTTGGC
 CTA CTGCCATCCGCTTAAGATGTGACATGCTCCTCCTTCTTACGCCATGATTGTGAGG
 CCTCCCAACCATGTGGAACTGCAAGTCCAATAAAC

SEQ.ID.NO. 397 HG1014938N0 PLI00014330E04.contig
 GGACTAAAGATCTCTCAGGGGAAGTAAACACATTAATAAAGACAGTCTTTAAAGGAAG
 TGTTCCTCCCAACCTTTCAGCAGTAAAGCTGTCTGAACAGGGCTAAGAAATCCTTCACATG
 AAAACAGCTGAGGAGCAGAGGCTCCAGGAGAAGTTTTCAGTGATTGATGGAACGTAAACC
 TTGAAAGTCTACCTCGTCAGTTCTTGAGATTGGGGTTCTTCAGGGCTTCATCCCGGCTTAT
 TTCCCTGTGGAAATCAGAGATAGAAAACAATGAAGCATCAGGTTGAAAACAGAAAGTGCA
 AAACAGTCAAGGAACACAAACCCATACATCACACCGAGCCGGAGGAAGCCCCCTGCC
 CACCAGAGATGACCAGGCAACAAGCCCTGGCCAGGCAACTCTCACACCTTTCCCAAGCTTC
 GTGGATAATTAAAGAAACAGTGGTGACCTAGTACCTAGGAACATAATCAGAACGTAGAGTAA
 TAACTGTTTTTCCAAACCACGGATCTTGGTAGGAGCTTCTCAGAGGTGATCCTTGTCAACCC
 GTCAGAAATGGTTGCCTCTCATTTATCAAGAAACAGAGGAGAAACAAACGACCAGTGAA

CTCACCATCCACTCAGGCATGTGAAAGCTGAGGTGCCCTCATTACCGTCCCACCTTCTCC
ATAGTGCAGGCTTGCGTAAGATATGTCTCTGAAGCAAAAGGAAAGCTGGTTTGGAAATAC
TGGGGTTATCTGAGTGCCCCAAAGCAGACCCCAACCCTAGTGAGATTCAGAAAGTTCACAGT
GGCTAGCAGAAGCTGCATCCCCTTTTTCACAGCAGCCTCCGATTCAATCAAGAGCATTAGT
CAACTGCCCTGACAAACCCCTTACTGGCCGGCTTCAAAGCCTTCTGTGTCTGATTTCAAAAT
CAGAGCTGATGCAAAGTGTGCAACCCAGGCACTGTGCTTCTCACAGGGCACTCACTGTTT
TAATGCTACCCAAACACTCTACCCGCACTCCCCCAGTGGGGACTTTTCTGCAGCAATAGTT
GGACTTCCTTCTTCCCTGTGTGTGAGGCACTTGTGAATTCTCTCCACCCCTTCAGGACA
CGTCACTGGAGGTGCGGATACCAAGATCAATAAAGACTGAGGTCAAGGCACTCCCAT
TTCTCAAGAGTTGCTTGATTTGGATCAGTAAACAGGCGTTTCCTTGAAATCAAGAGCTCC
TAGGGCTTTCAAGTTTAAACAGCTATGCAGAAAGCTCCCCATCTGCTCACACCAAGGGTCT
GGCACCAAACTCTCGCATATCTCAAAGTTAGAGTCTTATTGTCATGGGAGGTTTCGGATGGAG
CCTAGGAAGATCTTTTTCAAAATCAGAAATTTCTTAGCAGCTCTAAGAAAGTTCTCTGGGACG
ACAGTCCCTCTGCTGGCTCCCTGTGCTACATACAGAAAGGCCTCAGGGCCCCGACGTAC
AGACACACAGGTGCTGATTGCAAAGTGACTTCACTGAGGCTAGTCAGAAAGCCACACAGT
GTAGGTTTAGAGGACAGCCAGGGTTTCCAACCAAGACTAGTGTGCTGCCTCAAAGTCCCTAA
CCAATGCTGGATTTTACACACAGATCTGGTTTCAGGAAGAAACAGGATGCCAGTGCAGTA
GCCCTGTACACTATCACCTTAAAGGACCTCCGAAAGTTGTTTCTGCTTTATCTCTCCATAC
CAACTCCCTCCTTGCTGCACACTTCAGAAATTTTGGAAATAAGACATAACTAGAAAAAAGG
TAAGCTAGCCTATCTCACCTGCTTAGTTCATTCCTAGAGTATATTATCCAAATAAGAGTAC
CACAGGTTTCAGGCACAAAGGAAGCCTAACAAAGATATGCCTGCTTTTGGTCCCTTTAAGGA
CCAGGACTCTTTACAGTGTTAATTCCCAATTGAAGAGGATAGTTTCTCTACAGCTAACAGT

GACAATATTATAATTATGATTATAGGCTCTGTGAAAGGAAAAACATCTTTGGGCCCCCAAAA
TCACTAAGCCAAAGGGAAAAGTCAAGCTGGAACCTGCTTAGGGCAAACCTGCCTCCGATTC
TATTCAAATCACCCCTCTGCTCACTGAGATAAATGCACATCTGATTGCCCTACTTTGAAG
AGGCCAATCAAAACTCAAAAGAATGCAACCATTTGTCTCTTAATCTACCCATGACCTGGA
AGCCCCCTCCCTGCTTCCAGATGTCTCGCCTTTCCAAACCTGAAACCAGTGTTCATCATACA
TATGTTGATTGATGTCTCATGTCCCCTTAAATGTATAAAACCAAACTGTGTTCTTGACCA
CCTTGGGCACATGTCAATCAGGACCTCCTGAGGCTGTGTCAATGATGCACATCCTCAATCT
TGGCAAAATAAACTTTCTATAAACTTTTTTTTTTTTTTTTGAGACAGAGTTTGTCTAACA
AGAGAGAAAACCTCTGCACCAAAAAAAAAGAAAGAAAGAAATGAGGAGGTAAAGGCC
TAAACATGAAAGGCCTTATAAGCTAAGGAATTTGAATTTATTCTATGTATAATGGGCAT
TATT

SEQ.ID.NO. 398 HG1014939N0 PLT00014330E14.contig
GAGGAGGCAGGGAGTGAAACTGCCTGAGGAACAGAGAGAGGGGAGGAGAGAGAGGA
GAGGAGAAGAGAGCAGAGAGATCAGATCTGTAGCTAGTCACTGTGACAGGACCC
AAGTCCAAAGGTAAATCGAGAACAAAGTGATGGAACTAAGAATGAAGGAGAAATGGGTGGA
CAGGGCACAAATTGGGGAATGACAGCCTATTCACTGAGGATAGGAAATGCACATAAAGGA
CAGATGAAGATGATTTCTCTGTGTGATTAGAAGGATAGCAATTACAAAAGATAAGAGTG
AAAGGAGAGGGAGGAAAAGGGGAAAAGGGAATAAGTAAGAAAGACAAAAATAAGCAATGAGTG
TCTAAGTACCGATCTGAATGCTTAGAAAAATCACCAAGGCTGTTAGGCTGAGCCAAAAGAAG
AGCCACAGCCGTACCCAAAGAGGGGACCATCGGAAAAGACAGAAAGCAGGCAATTTAAGTAGC
TAGCTACTGGGGTAAAATGAAAAAAAAGAAAAAATAGAAGATGGAAAAGGTGAGAAAGCAA

AAGAGGAAAAACCCAAAGGAGACACTGGCCCAGAAACCTTTTAACTTTCTGTAGAATC
 TTCAAGGCCCCCAAACTAGGCCAAGATTCTGACTGACATTTCTGAGCTCACCCAGCTGCCGT
 TTCCTTGAGAGAAATTCTGACAGGAAGGCATAGAAATGCCAAACCACTGCCAGCTCTGG
 AAAATAAATAAAGTAAATACTATCCACATCCTAAACACCTCAGTGGGGGCCGAAGC
 CTGCTGGGGAAGAAAGCCCCAGCGAGGAGGCCGCCCTGCAACTTTTCCCCAGGCTT
 AGACAAATGATGCTTGAGACACACCAAGATACCTGCACACTAGAGGCTGCAGTCACCTCA
 GAAGAAAGATGCCCCCTGCTCTGGCTGGTCAACCGGAACCAAGTCCGTCCTCCTGAGAGG
 TTTGGTCCCCTTCAACCCAGCTACAGCAGGGCTGGCAATGCCCAGTCCTTGGAGAAACAGA
 AGAGATTCAAACTGCAACTGAAATTACCTACTAAACCTTTCCTTGAGAAACAATTCTGGT
 TTTTTCACCATATTCTTATGCTCCCAGGAAGGAAGAGAGAGGAGTCCCTGGAGTCAAAAG
 TGAGCAAGATTTTAAATGCTCAACACACACTGTTCTGTGCCCTTTGCACTCACAGCTGCAGC
 CACTTACATGCTTTTGACAGACACACAGTCAAGAAAAGAGTTTGAGGATGGAGTTTGTGCC
 AACCCTAAGCCAGCCAGGGGAAGAGCATGAGACATAGAAAGACAGGGCTCCTGCTCACTAG
 GGGCTAACAGTCTAAGGAAAGAAATTAGACATGTATATATGTTGTGATCCAAATTGAATTG
 CAAGTTATAGCATGGCTAGTTCCAAAGGCATAATTCTTACCTATCCCACCAAAATATGCCCTA
 GCACATTATGTATCATCAAAGTCTCTATTGATCTATTTCAGCCCCAAGATATCTATTCACTAA
 GCAATCATTTTCTCAGTTCCTGATCTGTAGAATAGGATAATAGTACCTGTCAACACCAGG
 CACAGTGGCTTATGCCTAAAACTCAGCATTTTGGGGTGGCTAAGACAGGGGATTGCTTGA
 GGCTGGTCTTGAACTCCTGACCTCAAGTGATCCGCCCGCTTGGCCTCCCAAAAGTGTGG
 GATTACAAATGTGAGCCACCGTGCCCGCCAGCCTATAGGGTATTTTTTATATTCCAACTATTAA
 TTCTACTAGCATGTAGCATACATGCATTCACCCCTAGTACATATGACTAGCATATAATATG
 CTTTACTTTTAAAAAATGCTTTTATACATGAGACTGATCTTACCATCTTCAATATCCTGTC

TGAAAGCCTTCATTTGTGAGTGCTGTGAGATCACATGTCCTCCCAAGTAGGACATGGA
 GTGTGAAGGCCATTGAGTGGGTGTTCCAAAGGGCAGCACAGAAATGGAGTTTAGTTGAAGG
 TGCTGGGGACACACTACCTGGGCTCAAGTCCCAGCTCCCCACTCTCCAGCTGTGCCACTT
 GAGCAAGTTGCTTGACCTCTCCAGGCCTGGGTTTCCTCATCTGTAAAGTAGGAATAATTGA
 TAATAATATCTACTTCATAGAGTTTTCATGAGAAATACATGTAAGGCCCTGAAACCCAGTATT
 GGGTACATAATAATACATTCTTAAGAGTCTTTATTCTTGACAATGTACCTGAAATGTTCC
 TGAAATTAAAG

SEQ.ID.NO. 399 HG1014941N0 PLT00014330E24.contig

TGCTTGCTTCTCTCCTCAGAGAAAGTTTGACATCTCCCCAGCACAGATATTTGGACGGAG
 CAGGCTCCTGAGGTCTTCAGGATACCCATCCAGCCTCCTGGCCCCACCCCTCCTGGTCCT
 TTCCTGCTGCTGATTCTGCTGCGGGGATGCACAGGTGAGCACTGCCCTTGGGCTAGAACCC
 TTTCTCTCTGCTTGCAAGCCCTGAAATCCCTGCCCTGGGCTTTCAGGGATTCCAGACAAA
 CACAAGTGCTCCCAAGGTAGGTCAAGGACTGGGCATTGTCTGAGAGCAGAAATCAGGGA
 GCTGAGCACTTCTCTGTGGATCTGCAGCTGGCTACCAAGGATGGAATCAGGTGATCCT
 TCAACCAGCACCATAAATTAGTTACATTTTAAAAATAATATTTCATTTTGACAAAACACAGAA
 GAATCATTAATTTAAATCCTTACCAGAAAGCAGACCTGGTTCCTGTCAAAATATTGAGCAGTA
 GGTTTTGGGAAACCCCTTCTGCTGCCCCAAACTGAGGAGCAGGAGCCTTGAGGCTGGAGCCT
 TGAGGCTCCAGGAGGCAAAACCATCTGCCCAAGGCCACAGAGTGAGCTGAGAGACAGCCT
 GACAGGTGTGTGGCTCCAGGCTCAGCCCTCTGTTTCAGCACCCCTGAGGATGTGAAGTCTG
 AAACAGTGCCTGAGTTGGGCAGGGGAGCAATGTTTCCCGGGGATGTTTCAGCCCACTGTCCT
 TCTTCAGGATGCTGCATACITTTGCCCTCTGTTCGCCCTGCTTCCCCACTCCTCTGTAATCT

CAGTGTTACCACCCAATTTTGTGCCCCTCTTCTTTAAGAAGGAATTTGCTGCTTTC
TGGGCCCTGTAAATCCTTTCCCTAGGGCTTGCAATAACAAGTACTACAGACTGGGCAGCTAA
AATTACAGACATTTACTTCCCTTGCACTTTTGGAGGCTGAATGCAAGATCAAGGTGTTGGC
AGGGTTAGTCTGTACTGAGGTCCTTTCTTACTGGCTCGTAGATGGCCGTCTCTGCTCTGTG
TCCTTATTACGTGTTCTTCCCTCAGTGTGTCTGTCTGCTCATCTCCTCTTCTTCTGAGA
TCGCCAGTCATAATTAGATTAAAGACCCCCCAATGACCTCATGTAACTAAATAAGCTTCTTA
AAGACCCCTGTCTCCAAATACAGTCAAAATCTGAGGTTCTGGGGACTGGGGCTTCAACACA
TGCAATTTTGGGGACACATTCAGCCCCATAACAAGTACCACAATCTACCTTGGTTCTTTTCG
ACTTTACCATGGAGAAATTTATTTTCCCTTCCCCAGGACAGGCAGGATTGGAAAGACCTTTC
CTGGTGCAGCCTGGGAGAGCCCTGCTCACAGCCCCACAGACAGGCTGCCTGTCTCACAACC
CTACTCAGGCCAGGAGCTGACCCCAAGATCACAGGGCTGGGCACAGCGCTCCCTGCT
CCCCACTGGCCCTGCTGTGAGTCACTCACAGGTCCTAGAGGACAAGGTTCTCACTTTT
CAAAATTAATCTTCTTACATATACACTTTTACCCAAATGACAGTGGTAAGAGTTCAGCAATT
GAGCTCTTCCTACAAAGGTAGACACAGTTCTTAGTAAATTTGTGTATTTTAACTCTACTTAA
CTCTACTAAATCCTCACAAACATCCCTATTTTGAAAGCTACACCTATTAACCTCCTATTTGAAA
ACAAAGGAGTCAGGCAGAGGGAGGAGCTTGTCCATTGTCCCCCTACCTCAGTCAGTGGAGCA
GCTGGGATTCAAGTCCAGGTGGACCAACTCCAGAGTTCAATCTTTTAGCCATGTTATGCT
ACTTGTAGAAAATATTAGCACATGCAGGGGAAGTAAACATAGGAATAATCATCACCCCATCAC
ACCAAGTCAGAGAAAGTGACACTCAGAGTTAAACGCTGGTGTATATACTCCCAACATTTGT
TACAAAATGTGTTGGATGACTTTTATGTGCCAAATATTTTACTGAGATTTTAAACAGGTTTC
ATAATTTTATTTAATGCACAAAACATCTTTATATAGTATAAGGACAGACAAACTTTTAT
ATATAATTTGTGAATGCATGTGTATATATGTATACATATATATATGTGTGTATAAGTGACT

ATAAGTGACATACATGACATAAGATTTCACATTTATTGTTATCTTTTATTGCAATG
 AAATGTCTCAGAAAAGAAATAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAA
 AAAAGAAAAAAGAAAAAGAAAAAGAAAAA

SEQ.ID.NO. 400 HG1014943N0 PLT00014330F01.contig
 GCTGCCATGTCCTGAAGATACTAAGCAACCGTGTGCAGAGGAACTAAATACCGAGTTC
 AACAAACAGGTCTCTTGCTATCCTTCAAGCAAAGCTATCTCCAAATCTGAAGAAGACCA
 AGAAATAGAAAAAGTAGCACCTCTTGCCCTGCTTCCCTGTCTAATAATTTGAAAAGATGGTG
 AGTATTGATGTTATTAATTGCCCTTTTCCTTTTATATTGCTACCAAAGCCATGAGTATT
 CTCTTTTTTCTTCTCTTTTAAAAAACAATAAATACTAACTCAGTTTTTAAAGGGGT
 TTATAACCCCTGGGTCTTGGTCCCATCCAAAGCATCTTCATTTTATTCACTCGGAAAA
 TCTGACTTTATGTGGTGAGCCCCAAGAAATTTGGGTTGAGGTGGTCTGAGGATTTGGGTT
 GAGGCCAAAAGCCAGTACCAGGGCTGAATAGTGGCAAAAATTGGGGCCAAAACCTATAGC
 AAGTGCAGAAATGATGGGAACTCCCAGGTGCTGTTCTTTTGTTCCTCCAGGAGGCTC
 TGTAGCACCATTTGAACCTTCTTGTCTTTTAGAATCATCCCCAGCTTTTTCAGTATTCTCA
 GACACAGCAGGAAAAACAGCCACGCAGCTTGAACTTAGGGTGCACACTGTGTGCTCACA
 TAGAGCTGGGAACACTCTTGGCATGTGTGCACACACCTTTTAATGATGCAGCCGTCATCC
 ACCAAGTGCCAAATGACGACAGGCAACCAGCTGTCTATAGCTTACATTGCCAGGGCCTTGTA
 GAAAAGGTATGACACAGAAAAACATCAGTATCCAGGAAGTGATTACGTCCATTTTCTGTGA
 CTCTGGGCCAGATTATGCAGTCTCTGTGTTCTTTCATGCCCCCTTGTGTGATTCTCCAGATT
 TGGACTGGGGCCCTCGTGGTCTGTGGTAGGTGATACTGGGGGAAAGAAACAATGGTCAC
 ATTGTATAAGCTATCAGTTACTGGCCTGGTCAAAACTCTCGGAGAAATAAGGCTGAAATCA

TGGGATAAAATGAAATTTGTGGATATACAAATATAATTTGTCTTTATGACAGTACTCAG
AAATCCACAAGGAAAATGTTGTATGGGAGTTGGGCATTGTCAAGTGTTTAGGTAGGAGAGC
CTTTATTAGGGTGACTGGATTTTCTTTTCTACCTTTTTTCTTTTGGTGATATAATATACGTAG
CATAAAATTTACCATCTTAATACTATTTCAAGTGCACCGTTCTTGGGCGTTCAGTACATTTA
CATTGTAGTGCAACCATCACCCAGCATCCACCCTCCAGAAATTCCTTTTCATCTCACAAAGTTG
AAACTCCATACCCATTAAACAGCTGTCCATTATCTCCTCTCCCGCTCCCTGGAAACCCACC
ATTCTACTTTCTGTCTCTATGAATTTGACTACTCTAAGGACCTCATGTAGTGGAAATTGT
ACAGTATTTGCCCTTTCGGTGACTGGCTTATTTCACTTAGCAAAATGCCTTCAAGGTTTCAT
CCAAAGTTATAGCATGTGCCAGGATTTTCTTCCCTTTTAAAGGCTGAATAATAATTTCAATTGG
ATGTCTCTATCACATTTTATTTATCCATTCACTCATTTGATGGAATCTTGGATTGCCCTC
CTTTTGCCTATGAACATGGATATATAATTTCTACCATTTTTCATGAGTTAAATAATCAGGC
TTGTAATTTTACACATTTTTTTTCTAATTTACTTTTGGATCTTTTGTATTTCAAAGGAC
CTAGGGTATGGTCTGATATGATAGAATTGACAGTGGGAATGAATCAGAGGCAAGGAGCAT
ATAAAGTCAGTAGAGCCACCATAACATTGCTAAATTTGGGTTCAAAGTCCTGGACTACTATA
GATAAGAAAATTACTTATCAATGATCTTTTCAATAATTCCTAAAAAAACCCCTCCCTTGA
ATGAAAATCCAAATTTGGTTTATACTGAGTTCTCATTAGAAAATAACATTTAAAGAAAAA
GTGACATTAAATTTGAAAGAGCACTGTTATAGAAACATTTAGAAATTTCTTATATGTAAC
AACTTTCTTTTCTTCTTAAAAAATTTTTTTTTTGAGATGGAGTGTCACTCTTTTGGCCAG
GCTGGAGTGCAGTGTGCGATCTCAGCTCACTGCAACCTCCACCTCCCGAGTTCAAGTGA
TTCTCCTGCCTCATCCTCCAGGTAGCTGGGATTACAGGTATGGGCCACCATGCCTGGCT
AATTATTGTATTTTATGATAGAGATGTGGTTTGGCCATTTTGGCTAGGCTGGTCTTGAAC
CCTGACCTTAAAGCAATCCACCCACCTGGCTTCCCAAAGTGTGGGATTACAGGCGTGAGC

CACTGCATCTGGCCAATAACTTTACTTTCTATCAGTCTGGAAATATTAATAAATATTGT
CAACTTGCAATTTGTACAAAAATCACTGACATATTTGGAAAGTGTCAATTCAGCAGGTG
TGTGTAGCGATGGCCAAAGAGGAATCGCAGAGCTGTGTCTCAATATGGTGGCCACTAGCC
ATATATGACTAATGAGACTTGACAAAGTGGTATTCTGAAATGAGATGTGCTGTATGTGTA
AAATACACCTT

SEQ.ID.NO. 401 HG1014944N0 PLT00014330F03.contig
AACGGCGCAGCTGGACTCTGTGTGTGGCAGCCTAAGCGAACCTCTTCACGGGCTAATG
ACAAATGCAGGCATTGTTGGGCTGTCTCCTAAGGTGAATCAACATAAAGAAAGCCCTGG
CTGCTCGTGCATGATAACTTTCATCTTTTCGGACATCTGGGAACTACCGCCACAAATGTCT
GGAAAAAATGACTTCAAACTAATGCAGCAGAGGCAAAACCATGGTTTGGCCAAAAATGG
AACAGAAATTTACAAAGGGAAGGGAAGTTCTCTGGAGTAATATGTGTAGGTCAGATGAGGA
TGGAACCTGGTGAAATGAGCAGAAGAGAAGTTGGAAAGAGTGCAGTGGTGATGACAGCCCTC
CGCTGCTCCTTCCCCAGAGGAAGGAGCTCCAAGAGATTCTCAACTGGCTGAAGGATACAC
AAGATGTCTTATTGTGTGTCACTTGTGTTTGTGTTACTAATGACTCCTGCTGACACCACCC
TTGCTTCTTCAGGCAACAAATGCTGTTTGCAGCACTTGTGAAACCCATTATTTTAAACAT
GAAATAAGCTTGTTCCTGACAGATTGCAGCAAGAAAGGAGGACGTTTGTCTGAAACGTGAT
CTAAACTCACTTCAAGAAAGAAAGTACATGCCGGGAAATTGAAACCAAGAGCTTCTCTTCC
ATCCTGGAGGCCCTTGACCTCATGGAAGGTGACTCCTTTGAGCTTTGTGTCAGGTCAGC
CATAGCCTATGCAGTGTGCAAGGTGACACCTTCCTTAGGGGCTAGCGTCCCTTTCTGCTC
CTTTCCAGCCTGGACGTCCCCCTGCTGTTCGCTTCGGTCTTTTTCGGCCTCCTCCCTGAT
GTGGCTTTGTCTCTCTGCTTAAATGCCACAGGCTTTTGGTTCTCTGCCACCAAAATGAGC

SEQ.ID.NO. 402 HG1014946N0 PLT00014330F04.contig

TTTGGATGTGAGTCCACTGAGCTGGCGTGTACCTACAATGAACAATCCTTCTGTACTCTC
ATATCAGTCTCTCTGCTCCTCAGTTTCCTGCAACATTTTGGTGACCAACAAGGACCAG
AGACAGCAGGTTTACTGTTTCCCTTGGCTCTGGCGCTGGAGCCCTGGCCCTCAGGAGAC
CTGTGACCCCAAGGTACCACTGGAAGAACTTCAAGCCCGGAGAGGAATCAGCTCTCCCCCTG
ACCCGGCGCCCTACCCAGCAGTGCAATGGAACCTGAAAGGAGCTACAGGACAATTTTCAG
AAACAGCACCCCTTCAGGAACCTGCGTTGGCTGAAACCGGCAGCTCAGGACAAGTGAACCCAG
CCAGCAGGACCCAGACCATTTGACTCGGCTGATCTTTGTGACAAAGACTGAGTTGCTGCTGA
GTACAACAGCCCTGCTCTGGTCACTCCAGAGGCTGACCCAGTCTACACACGGGTGAAGCCTT

GAGGAAACAAGCTCTGCTCTAGTCACACACCAGAGCTGACTAGTCTATGCACAGCT
GAAGCTTGAGGACTTGTC AAGCAAGTA AATGTAGTTAGAAATCTTAGGACTAGTAGTTTT
CCTGTAATACTGTTTTACTATTGTTCCGTCGCTGTGCTCAATCTTCTTCCCCAGGTAAGG
ACCTCTTTTGTCTCTGCTGGATATGAATATGCTGTACATTGCTTTGCTGTGTGTTACCCCC
CTTAACCATGCTAGAAGAATCACCCACAGAAAGGGTGTCTCCACCGTACACATACTACTTG
GTCAGGGAACAGCATAACTAAACTCTATTGTACCATACTTATTATGGGTGTATGGGAA
TCGCTTAGGAACTTGTAATTATAGTCA AACCACCTACTCAGTTTGTGACCCAGGAAATAA
CCAACTTTATGTATGTTATGACCCCAAGTTTTACCTGGTGAAATGGTTTGAAATTCGTGC
AGTCAAAAGAAAGTCTCTCTTAAACCAACCAGGGTCCCTCCCTTTTACTGAGGGACTA
TTTCTCTGTATTTTGATGCTCGTTAGGTAGCATACCAAGACTCACCCCATCTGTGGTAACC
TACCTTTGGAAAGATACTATAGAAATAACTACAAATATGTATGCACGCCAAGAGCAACTC
CACATTGTTCCCTGGGACCCAGAGAAAGATTGTTGGTATTGCATATCATTGGCACCCCTA
GCCACAACAAAGGTCAATTATGCTTACCAAAATGCTGGTGGACTGCAGGTTCCCTCCTTAC
ATGTCTACATACACACATACCCAGCAACCTTCATATATGTTATCAAAAAGGGAAACCCAGA
CCCGTCTGGCCAGCAAGGCACAAAATGAGGAACGCTATCTATCAGAACAGATTAGCTT
TAGACTATCTCCTAGCCCTAGAAAGGAGTAGTATGTGGAAAGTTCAATTTAACAAATTGCT
GCCTGGAAATCGATGACAATGGAAAGGCAATTATGGAAATAACTGCAAGAATGAGAAAAT
TAGCCCATGTTCCAGTTCAAACTTGGAAAGGGTGGTCTCCAGATTCTCTCTTTTGGAGGCT
TGGTTTTATTTTCTGAGGGTTCAAGACTTTAATAGGAGTGGTTCTGGCCATATTAGGA
AGTTGCCTAATACTCCCTTGTCTCTTACCCTCCTTGTTAGAAGCAATTCATCGAGTATA
GAGGCAATAGTAGCTAGGCAAAACCTACCCTCAGCTAATGGCTCTGCATAAATATCAACCT
TTGTCTAAATAAGAAACTCGTCCCTTCAGGCAGAACTAAGTAATAGTGATGCCTTCTAT

TAAACTTCTTTTATAAAGGCATCAAAAGGGGGAAACTGAGGCAGAAATTTAAAAATAAA
TATGCATTTCATTCACCTCCAAGAAAGTAACAGAAAGGCAAGGGTTAAAGAAAGAAC
AAGTTTTCCTCTGCCTAGCAGAGGAAAGTGCAGCAAGCTCACTTCAAGGACAGTAATAAG
ATAATACTGTTTGAGAAACCAAGGCCAAAGGAAATGGACTCCAGATACCCCCCACACCCAG
AGCAAGGTTGAAGGAAAAAGAAAGACAAATTCCTTTACTGTTACTCCTTTCCCGATC
TTCTTAAGCATGATTATGTTTTTACAAATGTCTGTATTTAGCCGGTTGTTTTCCTTC
AATGCAGCTACAAGGCCACCACCTATGCAAGGCCACAAAGTTATGCTATAGCTATAGATTAT
ATGACCTATCATTTGGTTAACTGCTTTTGTATTTTATTTATTTAGTCCACTTATAAAAC
CCGATTCTGCTTTGTTTAAATGCTCAGCTTTGTGGATGTGAGTCCGCTGAGCCAGTGTGT
ACCTAAATA

SEQ.ID.NO. 403 HG1014948N0 PLT00014330F05.contig

AAAATTTATCCCCCAGGGTTGAAAGTAAGTACCACAACACTAGTAGCTTTTAAAAAGCTTTT
TGTTACATGTTTCCTTTGAATTCTCTCTATATATTCAAAATAACATTATTATTATTAT
TATTTTTCATAAAGGCAACTATAATTGCTGTTTTCATTAAACCTGCCAGCTATGTTTTC
TTGTCCGGTGAAAGTTGTGTCAGTTTAAATCACATGCATTCTTTCCACAGGAAAAATAA
ATAAAAGTTTAAAAGTTTTCCTAAAGTGATTACTTAAATTGTTAAATAAAAGTAATAAGAA
AAAAATCCAATCAATATAATAACGTATGTATTAGTCAATTATAGCACTGCTATAAAGA
AATACTTGTGGCCAGGCATGGTGGCTCATGTCTGTAATCCAAACACTTTGGGAGGCCGAT
GCGGGCAGATCACTTGAGGTCAGGAGTTCAAAGACCAGCTCGACCACACATGGCAAAACCCC
GTCTCTACTAAAAATATGAAAAAGAAAAAATTAGGCTGGTGTGCTGGCACATGCCTGTAG
TCTCAGCTAGGCTAAGGCATGAGAAATTGCTTGAACCTGGGAGGTAGAGGTTGTAGTGAGC

AGAGATCATGCCATTGTACTCCCTTCTGGCGCAGAGTGAGATTGTGTCCAAAGAAAGAA
GGAGGAGAGGAGAAAGCCCAAGGAGAGAGAGAGGGAAGGGAAGGGAAGGGAAGGGA
CACTTGAAACTGGGTAATTTATAAAGAAAAGAGATTTAGTTGGCTTATGGTTGTGCAATC
TATACAGGCCTCAGGAAACTTACAATCATTTGTGGAGGGGAAGCAGGCACATCTTACATG
GCCACAGCAGGAGGAAGAGAGAGAAAGGGGAGGTGCTACACACTTTTAAACAAACCAGATCT
TATGAGAACTCACTAACTAACAGGCGAACAGCAAAGGGGAAATCTGCCCCAATGATCCAA
TTACCTCCAACATTAGGGATTACAAATTTGACATGAGATTTGGATGGGGACACAAATCCAA
ACCATATTAAATGTATAAAATTTTAAAGACAAAACTTAAATAATTGAAGGTACTTAATAA
TTTTACAAACTCATAATAATTGAGAAAGAAAGCTTGGGCTCATTTACACATTTTAAATAATCT
GGCTGGAATAAATAAATAAGTTAACATTTTAAATGTTAATTACCTCTATATTGGCAT
CTTGAAATAAGTATAATTTAAACAACTGACTTGCTTATTCATTTTCAATTGCTTT
GTTCTGTTGTAGACACTAAATTTTGTAAAGAAAAATATTTTAAATATATTAAATGCAAT
GAGGTAGCCCTCATATACTACTGATGGGAATGTCGAATCCTTTGATTTATCATATCAACT
TTTTAAGAAATCATTGTGGAGCTTTTAAATCCAGACCAAATGAAAAAAGTGCAATCCACCG
CATTCCTCTTGCTGACTGTGGCTAAAAATCTTGACAAAAATACATAAGGCAAATCTGAGGA
GACTCAAGGAGGTGAAGAGAAAGGATGCCAACTGGCCTCAGGACTTGAGCAATTAGCAAGT
AGTGAGTTTCCTGGGATTTTATGCTCTCAGTATATCTCAGTCTGGGTGCTAGAAAA
GCTTGCCGCCCCAGAAACCCTAAAGGATGTGCAGCCAGTGCAACAAGA

SEQ.ID.NO. 404 HG1014949N0 PLT00014330F13.contig
CCAGTCGCCCCGAGGCAGATGGGAGAAAGACTAGCCACAGCTGAGGACTGCTGCCCCGAGCC
ACCCTTGCCCTCCCTTCCCTTTCAGTCCCTCACCGGTGCGCGACTGCCTCCTCAGACAC

CGCGCGTTCCACTCAGCCGGTGCCGACCCCGCCTCCCGACTGCGGGA CTGGGGCG
 GGGAGCGAGGAAGGGCTAACGGCGCGCGCGCCTCCGAGTGGGGGAGCGGGAGGGCG
 CCCGAGGTCCGGCCCTCGGGATGGGAGGTCA CGCTCCTGTCCCTCTCGCCCGCGTCG
 GCTAGTGTCAGTCTGGCGGGGCA GTGTACCGGTTGCTGGGAAGACTGGAGGCCAGC
 GAGCAGCGTACCAACCACTCAGCCTCTCCGGCCCCAGCTGGAGCCGGGAAGCAGGCGGGG
 CTAAAGAGAAAGCCTGCGGGGGACTCCACTCCGCGCTCCCGCGGGGCCAAGGCCCGG
 CCCTCTCTTCTGGCTTAACCAATTCTGCTCCGCTTTGTGTGTTGGCCCA CCCCCTGGG
 CCTCGTCCCGCCAATCAGAGGAAAGGGCGCGCATCCCTGCGCGCGCCTCAGCCCCCAGC
 CGCGGAGTCCCGGAGCTAGCGGTCCA CCGTCTAGTCTTCCAGCAGCTCTTCCCCGCGGC
 CGCTGGCGGACACAGACAGGTGTACCCCTCGGTGCTTCCGGCGCGGTAGCCGGTGCCCCG
 GCGCGACGGGCGCGCAGCCATCCCCGTAGTTAAGTTAGCGCCTTCTTATTAA GTGA
 GGCCAGACTGGATAGCTATTATCTCTCAACCCTCACCCACAACCCTGCCAGGGGAAATATG
 AGCATTTATCCTGTATTTTAGGTAAGGACCCACACAAGTTAATGATACGCCCA AAGTCGCA
 ATGCCAGGGATAGGGGTTACTCCTGGTCGGCTTGCCAGCTGCCTCTCCATGAGCTTCCTG
 GTTGAACTGCCTGTCCAAATGTTGACCCCTAATGTCTTCTACTAGGAAATGCTACTAGTAT
 TCCTTTCCGGAAATGCATTTTGGGAACCTGGCCATCCCTTAGTGTGCCTATATCATCAGA
 ACAGACGCCCAGGATGGCCCTCAAGTCCTTTACTGCTGTAACTTCTCTTTGGAGGTCA
 CTTAGTACCAC TCTTCTGAGCTCTGCTCAGGGGGTCCCCCTAATTAGTGTA CTTCGACCTG
 CCATGCTTGGCCTGGAAACTCTTACCCTCTTCTCTGTTTAAATTGTAAACCTTAGTC
 AAATGCCATCTCCATAAACCTATGCTAGAA TGCCCCCTCTCTGTCTCAGTCCTGGGACA
 TATCTATGCTGCAGCGCAACCTCCCGCATGTTGATACTGTTGTTATAGTCAGATGAATAC
 TGTTCCCCCATAGCTTAAGGGACAGGCCTTGTCTACTCATGAACAATTGATCCCTTGAA

GCATGAGCAGGGCCAGCACGCCACAGATGATCAATGAGTGTGTTGCTGACCCGCATATG
 ACAACACGGATTAAAGGAAGACAAATCTGGTAGCAGCAGGCAAGAAAGCTGAAAAGCAAGA
 CACAGGATGCAGGTCTGGAGTGATCAGGGCCTGAACCAGGAAGGGATAACATTGAGAGAG
 ACATGGAAAAATGCGCTTGAAAGTATATATATTTTATTGTAAGAATCATACACAGATAAG
 GTAGTATTATATAGAGAAATGATTAGTTGAGGATATTTCCAGAGTCTTTAAGCCAGAGAAA
 TGTACTAATTTAGACATTTCTGGACTCATTTGCTCAGTACTGATGTCATGTTTAAAGGAC
 ACTGACACACTTGGCTAATCCCAGAGACATGAAGAGGATGGTGACTAGGGTGGTCACATA
 ATTAAGGGGCTGGGATATTTAACCTAGAAAGAGAAACAATTTTGTGCACATAAATGAC
 TGTACACAAATTCTAACAACTTTTCAATGTGAAAGAGGACTGGATTAGATGACTTCTTCT
 GTGTAGCTTTGGAGGGAAAACTCAACAATGGGGGAAAGAACAGGAAGCTGGATTTAG
 GTTTCCTATGTTATGGAAGTCATCACAGTGGGGATGATAATTGTAAACAGGAAGATGGAG
 GGTAGTGAAAGAAAGTTCAAGGATGAAGTTATGTGTATATTTG

SEQ.ID.NO. 405 HG1014950N0 PLT00014330G21.contig
 CCACCTGCACCCAGGTGAAATAAACAGTTTTATTGCTCACACAAGCCTGTTTGGTGGTC
 TCTTCACACAGACGCGAGTGAAATTTGGTGCCATGACTTGGATCAGGGGACCTCCCTTGG
 GAGATCAATCCCCTGTCTCCTGTTCTTTGCTCTGTGAGAAAGATCCACCTACGACCTCA
 GGTCTCTAGACCAACCAGCCCCAAGAAACATCTCACCAATTTCAATCCGTTTTC AACACAGC
 TCATAAGGAGAAAGGAGTTGGAATCTGCATCTCATCCCTTCCCAAAGGCCTCAGGCCTGAG
 AATTGTACTTCTCTCCAGAAACCAAGATGGTTCCACATCACCTTCTTCGAGGGACAGAC
 TCCTCTTCTTGACCTCCGTCTCCATGTTCCCTCATGACACTTCACTGTCCTTCTCCAGCA
 TGGACACTCTTTCCATCACTACCGCCCTGGAGAAAGCTAGGTCCTCTTCTGTGGAAAGTCACT

CCTTCCCAAAACCTCACACGAATTTTCTGAGGAGCCCCAGTGCCCCCTTCTCAGCCA
ACACACCACCTCCATTCTCTCTGAGCATCCAGTTTGGTCAGCACAGTTCTGCTCTCCC
CCTTCCAAAGCTGTTCCCAAGGTTATCATCCAGGGATATGTAAGTTTCCCATCTGAAATGCA
TCATTCCCTCTCTCTAGGCAGTCAAAATCTCAATTTCCAGGCCACATCATCTGATTAAAT
TCTTAAACATCAGAATAATCATTAACAAATGATACATCAGTTTCTAGATTCCATCTTAAC
CAAAACAACACACAAAAAATTACAAACAACCTCTCTGGGAAGAATTCTTTTGCCCAATAGG
ACCAAAATGCTTAAACTATTCTCTCTCCAGGATAATTTTATTTCATTCTTAACACTCAGCTC
AGGTGTCACTCCTCTGTGGAGCCTGCAGCAAATTCTCTTGAGACAGAGTTAGTCATTTCA
TTTTCTGGGCTTCACAAAGATCATTTGTAGCATTTAACACGTTGTATTATAATTACCTGCTA
GGAAGTCTGCCCTTTCCCACCTGGCCATGAACCTCCACAAGAGCGAGAACTGTGCTTATTC
ATCTTTCAATCCTCCATTTCAAATTCAATTACGCAACTATTCAACACCAAGCACTGTGTCA
AATTCAAAGTTAGTACTTAATAAATATTTTCAGGAGAGAAATGAAAGAGTGAAGGGAGATAA
GTTTACATTATCTTAAAAAATTCCTTTCAAAAAAGTTCTGGCAGAAAACTTCCTTTTTTATGT
CTACTTCCAAGGAAAAAGTGTTTTTCCCAAAATCCCATATTTAGATAGGGACCCCTCAAGATT
TGTTTTTCCAGTTTGTTTTATGTTTATAAACTTAAATTTAAGTGATAAATAATCTTGAAATT
ATATGGCATAGTACAAAGTGACCTTGGAGGCAATCTCATCCAAGGAGACAGCTAATATATG
GCACTCATGGCCTGCTCTCCACCTCCTGCATCCATAGCAGACAGCACGAATGGATCATGG
TGTACTTCCCTTATTAAGCCCCATACTCAGCCTAGAAATCTCAACTCAGCACTTAACAAGCT
GCTTTCTGGAGCTGGTAGTCAAGATGAATCCTGCTTGCCATCCTTGATCTAGTTTATTCT
CTCTTATTGCATATAAGCAGAGTGAGGTCCAGAAAGGTTAAATGAATATGCTTAGATTCC
AAAGCTAATTATTAGAGCCTGAACTAGACCCAGGTCACCTAAGCCTCTTTCCACCAGAAT
ACACTTTAATTTTATAGCTATAGGATTGAATACTTCATAGAAAATGTTAGTTGTGGTCAAT

CACGTTATATGTATATATTTATGTATATACAATGATGAAAAAACCGACATTT
TTACACTATATACATATGTATGTATATCTATGACAAAACCAATATTTTCTATGAATT
GTTTATTCCTATTAATCAGCTATAGAATATGTATGTTACATAATATATATAGCCACA
TAACATTTGTGTGTTATATATTACATTACATATTATATAATCTATGTTTTTCAGTCAA
AAGTGAGGTAGCTCTAAAAGTGTGCCACAAAAAACAGTAAACCCAAAAAATACAATG
AAAAACAAGCATTGGTCAATTTAACTCTTTTGATATTTTCTTATAGCTTCTTAGCATATG
CAAGACCTATTTGGAAAAAAATAGCTGTGGAAATTAGAAATTTAAGGCTGGTTATGTATA
TTATTTTAGTCAAGGGTTATGGAGTTAGTAACATAATACTAATTTTCACTGTGTAAATTCTG
CTCATCTCTCCATGTGATGACACTGCCACCAATACAAGGGGAATCAGAGAGGATCCC
GGAACGTGCTTAAATGTTACTGTTTAGGCAACTCCAATGCTAGAAAGTATCTCTGTTTATCA
CACACAAATGTGTAAAGAATGTATGTTAGAGAAAAACAAGAAAAATATTAAAGAAGTGAGTA
CAAAACCATTTGTTTTCTTTTTCAGAGCATGAGTTTAAATCTCTCCAAAGGGATAGGGTTGT
TTCTTCCAAGTGGTCAAGCAGGGGTATGAGAAAGAAACAGAAAAA

SEQ.ID.NO. 406 HG1014953N0 PLT00014330H06.contig
GAGCTCATGACCAGAGGATCTGTCTTATGCAAAAGGTCAGAGAGGCTTCCCTGATAAAGG
AAGGTCATCCAAGATGTGCAAAGAGTAGCAAAAAAATTCGGTCGCCAGGCATAAACATTT
CCCACAGAGACTGAACCTGGGCAGTCTGTCTTCTTGTTCAGCACTCATAAGGTGTTACA
CCACAGTGAACTTTTAGTGTGGCTTCCGTCTCAGCATCTGCTTCCCAGAGAAAATCAATG
AGAAATGATCTCTTATCCTATCAAGGATCATTTTACCTCCTTCTCACCAAATGTTACACAT
CAGTGAGCAGCACTGCCTAGGAGATAACAGGAAGCCGGGTTTTCACAGCAGCCATCCTAAC
CAACTAATTCTGTGATGATTGAGAGACTGTGAGTGACCCGATTGGCCAGACCCCTTACTGA

ACCCAACTGACAACTGTGTCCTGATTAAAGGGGAATGAAATTTGTTTTCACAAATCCCATA
TAGACCTATAAACTTGTCTGACCTGTCAAGAGTGCCAGTTTTCAAAACTGCCATTAGCAA
ACCACTGTGAAACCACCCACAAAAGCCTTACATTTTAGGCCTTTTCATTTTATGTTTACTA
TGAAACATTAAGAATGCATGAATATAGAAATGGAATTGGTCCCCTTTTCACCCGATCACA
AACTTTGTACCTCAATTCTTAAAGTCAGGGATGTTTAACTCAGCCAGCCACCTAGCTAGG
GTCTGTAAACAAACAGTTGGTTATTTTTACCCCTTTGTCTACTCAAAATTAAGAAAGTAAA
GACGATAAAACAATATTCTACAGAAAAAAAGACAAATGTTAAAAAAGTAATAAAAGAG
AACTTTCTTTCATATTTCATGTGAGAAGGATAGATGTTTAGAAATAAAGAATAAAAATTTGT
GTGAGGGATGCCTCTTGCAATTTTAAGAATTAGAAGAGGTGATTGGTAAAACTGTAGACAT
GGACGAGCTGTTTTCCAAAGGCACGGGATCAGCTACAAGTACATAGAAAAAGGTCCCCTCAT
ATAAAAACAAATTATGTCAAGGTGAATTTGTGTGTTATATATATGTACTAAATGATGTGTA
CATGCAACTGTTGTTACAGGAAAAAAGGAAAAACAGAAAAAACTGAAAAAAATTACATGA
AAAAAATGTATCAAGAACCCACAGTAGTTGTGTAATAGCATGATACTGTTATTGAGTACCT
GAAATCAATTGCAGATTGACATGAGGAGGAGCTCAATGTAACTTCTCCCAGAGATGATCA
CAGAACTTAAAGAGGCTGAATTTCATAGAACAGCTTCTGTGTAGAAAGTCACTCTATAGTTC
CTCCCAGAGGAGGGGACTGACTTTTCAATTTGTCCATTGGGCCATGCTCGGACCACAGTA
CTTGAAAACAGAACCTTAGGAAAAATGCATCATATTTTAATTCTGAACCAAGAGTTGGCTCA
TTTTTAAAGATACCATGCTTAGTTTTTCATTTATTTCAGATTTAGGTAAAGGGTTCAATATT
CAGGTCAGGATTTGAAATGGGCATCAGGAAGGTGCCCTTAGGTGAGGAAGATTTGTGATG
TGGGCAACTTCTTGCTAATAAAGCCTTCAGTTCAACCTGTACTGTGTTCAAGAACTTCTG
AAGTTTTACAGAGGCTTGCCTGGAAATCTGGGATGAGGGCCTGGGGTTGAGGTCAGGCAGC
TTCGTGAACGGAGAAGTGTTGTGCTAATGGGGTTGCTTGAGGGAGGACAAATGATATCCAC

CTTTTTC AATAATTTTGCCTAAGGCCACCCCTGGCCCTGCCATTCAA AATCACCTTTTCTG
 TCATTTTGGCGTTTGTGTACATTGCAGAACCCCTGACTTAGGCCAGGAGGCCATTTAGGA
 AATTATTCCTGACATTTTGTTTTAAATTAATAAAATACATAAAATAGGGTTCCTTGTAAAT
 AGATCTCACACATCAAGCATAAGTAGATCGCATTTATTTATTTTAAATGACATATGCTGT
 AGCTCAGTTTGTAGAATTCAAGAAATATAC TTCATCTTTTGTGTAGGTAATTTTCATTTCAA
 TTTCCACAGTCCCTCTTTCATTTTGGAGAAATAAAATTTGTGTTTATGTCTATATACAC
 TCTCCAATGGGGCAATTC AAAAATAAAAGTGCATCCTTAACACTGCAATTTCTTTGCTTTT
 GGCTGCACTAGAACTTTTCCCTCATTTCAATGTAGCTTTTGTATGGGTGAATGCAATAA
 AAAAGGTAATAAAC

SEQ.ID.NO. 407 HG1014954N0 PLT00014330H12.contig
 GGGCGGAGGCAAGTGCTGGAAGGTGCGCAGAGGGCCGGGCTGGGCTGGGAGGCGAGG
 CTCGCTTACCCGCCATGGGGCTGCTCCCGCAGGCTGGTCTCGCCGGTCTGGGCGGCTGGG
 GCTGGGCTGCCAGGCCGTCTCTTAAAGCGCCGCGGGGCGCGCTCGGGCGAGCACAGGGGC
 GGTCCGTGGTGTGCTGCCCCAGCTGCGGCGCGGGGAGCCGCGCAGTGCACGGCAGGG
 GCAGGGACCCAGCGCCCCGAGGCCGCAGTCCCCGCCCCAGCCCGTCCCTGGCCAGGCCCA
 GGCCCGGCGGTGGAGGGGCGCAGAGCCGACCGTGAGCTCCCGGAACCCGAGGGCTCCAG
 CCAGGGCGGTCCCGGGCGCCAGGGGCGCCAGGGGCTCCCTGCGGCTTGGGGCAGGGTGG
 GGTGAGAGTACCAACCCCATCCCGGAGAGCAGGGCTACAGCCTTCAACGCCCGCGTCTGAC
 ACAGGGTCCGCACCCGACCCCTCACACCTGGTGCCAGGACTGAGCCCCCGCCCTCCCA
 GATCTCTCTCAAGCCGAGGCCCCCTCCAGCCGCACTCCGACCCCGGCTCGGGGAGGCCTT

CCGGATCCCGGGGTGAGTCTCCGCGGGGTCCCCGGCCCCAGCGCTGCCTGATG
 GCCTTGGCAGGATGCGGCGCTCGGCACCCGTTCTGCTCTCAGCGGACAAACCGGTGGA
 GAGCGCAATTCTGGAGCGGGTTCAGAAAGCTCCTGGGTTCCAAATTCAGGCTCCTCACTCG
 CCAGCCCTGCGATTGTGGACAACCTAACGTTTCGTAAGCTGGGTTCACTCTCTCAATTGG
 GTTAATGTCCGTGAGGGGCTCTCGTAAGGGCGGCTGGCCCTGCCGTAACCTAAGAAAGCAT
 GTAAAGAGCACTTAGAGCGGAGGAGCGGCGGCCAGCCCCCTGTTACGGTGGTTATCGG
 ATGGAGCCAGCGTCGCCCTGGGTCCCCTCTGTACACCCAGCGCCTATTATTATGGTGATT
 ATTAGAACGCAGGGCCCGTGCCCCACTAGTCCAGGCTGGCGAACCGGCCCTTGCGGGAAAGG
 CGGGAGTAAGGTGCTGCGTGTCTCTCTGTGTACACCTCGCGCGCCCTGTGTCGCGCCC
 CTATGGTTCCCCCGCCGAGGCCCTCGAGGCGTCTGAGTGGGAGCGAAAGGAATGTGTGTA
 TCTGTGTGCAGGGGAGTGTGTGTGTGTGGGGGGGAGGGGTGGTTGTGGGGAGTG
 CGTGTTTTGGGGGATGGACCTTCTCGTTTTACAGCCCTGAAGGTTATGTCCAGAGTCAT
 AGCCTGGGTCCACACCCCCCGGTCACTGTGCTGTCTAGGGGCGCAGTCAGTGTGCGT
 GAGGCTCCACCCTGGGCTTTGGCTCACCTAAGGCGAGGAGGGTGGTGAACCTGGGGT
 GCCCACTGGGCCGGTCTGGGAGCCCTCAGCGCCAGGGTGCAGAGGCCCGCAGCCACTCC
 CCTTCTAGGCTCTGGGCGGTGGGGGCATGAGAAAGCGGAGACCCAGGCCACTTCACCCAG
 TAACTGTCTCAGCTGTGCGCTTAACCTGGCGAGATCTCCGCTCTACTTGGAGACAGTCTG
 GGGTGTCTGGAGAAATCCTCCCCCACTTCTGCTTACCCCAAGTCTGCTGTGGGAGCAGGTGT
 CTCAGGAAATGGGGCCTGGGTCCCACCCAAAGCCCTCGGGTCTGAGTTAGATGTCTGTATG
 TTTGTTGTTGCCCCCTGCAACAAAGTACACAGATTGAGTAGTTTAAAGCCACATGCAATGTT
 ATCTCACAGTTCGGGGGTCTGGGAATCTAAACTGGGTTGGCAGGGCCTGTGCCCTTCTGGA
 GGCAAGGAGCTGTTTCCCAGCGTTTCCAGCTCCTAGAGGCTGCTAGCTAGTGGCCCTTT

CCTTTGTCTTCAAAGCCGGCAGCACAGCCTTTTCTAACCTCTCCCACTCCAGGTCAITTC
 ACATCAACTCTTTTCTGCTCCCTCTTGTAAAGTGTATACCTACCTTGTTTTAAACCTGAA
 TTGACTGTCCCTTAGCTGAGAGAGCCAGACACACTCCATTTTGGCTCCTTCACTTGCAGC
 CCTTACCCACCCCTTCTCTCAAGGACTTAACTGTGCAAGCTGACTCCCAAGCACATCA
 AAGAATGCAATTAATGGATAAGATACTGTGGCAAGCTATATCCGCAGTTCCCAAGGAATTT
 GCCCAGTTGATAGTACCTTAAGCCCCCACAATTTGTGTCCGGTTGATGGTACCCAAAGCCC
 CCGCATCTATCGCCTTGTGATGGATTTAAACCCCTGCACCTGGAACTGTTTGTTTTCCCT
 GTAACCAATTTGTCTTTAACTTTTGCCTGTTTGTCTTCTGTAAAGATGTCTTCAGCTAGG
 CTCCTCCTCCCTTTTCTAAACCAAAGTATAAAAGAAATCTAGCCCTTTTCTTTCAGGGCCG
 AGAGAAATTTTGAGCACTAGCCGTCTCTCGGTTGCCGG

SEQ.ID.NO. 408 HG1014956N0 PLT00014330H14.contig
 GCGCAGGGAAGTACGACGCATACGGACCCCCGGAGCCCCGGAACCTCGCTTCCCACGCCA
 GACCTCATCCCCCTTCTGACAGAGACGCCCTCGTCCCACCGTCGCCGCGCAGGCCAGGGGT
 CCTCCCCACCGCCGAGGTCTAGCCGCCGCCGCCGCCCTTTCAGGGCAGCGCGGGACCC
 CCTGGCCCTCGCCCAAGTGTCCGGGGCTCACCCGCCGCTCCTCGCCGCCCTCCGGAGGCCA
 GGGATTGTTTCCCAACAAGGACTGGAGTCCAGGGAAGGACTCGCCTGGAGCGGGAGGCCG
 GGATGCACATGCGCAGAGGAGAGGCCGAGCGCCCGAGTTAGGCTGGTGGATGCGATGCG
 GGGCTTGAAAGCGGCGCTAGACGGTGCCTGTCTGAGCTCTGGCCGCAACCCCTCAGTTGTAAA
 ATGAGGCTGATAGTGCCCTACCTAGCAGGGGTTTTATGAGAAATTGGAGGCGCTTACGTTGT
 AAAAGCAATTAACCGGTGTCTGACACATAAAATTGTTCAGTATATTTCTCACAAATTATAT
 ATCATGCTGGTCATCAAAAGGCTCTCTCTGCAAAACAGCCCTTGTATAGTGGTTGAGAGAT

TACTGTCTCAGTTACCCACTCCTGCTTGGAACCCAGGATTCTGTCTCAGTGTCCCCAGAGGTT
GTGGATCAAAATCAATGTGCCCTGTTAGACACAAGCCAAGAGAAAGGTTGTGCTTGAAT
TCATTACGATCATGCATGGTAACGTTTAACTGAAAGACTGGAAAGAGAAAATAAAATCT
GTGACCAAGTATGGTTATTCTCTGGAATAAGGAAAAACAGAAAGCTAGCAATTCAAGACCCA
AACATAAAATTGTTTATATGTTTAAAAAAAAGCCGTTTGAGTAACAAATTTGAAGA
AGAAAGCAGCACTCATTTCCCAAACCAACCATACTTTACCCAGAACTACACTACACGT
ACATTCTTTTACGTGCACATCTGGCATTTTTCAGCACACTATGAGCATGTGCGTGTGTTAC
AACTTTTCATTTAAAAAATCAACTCCTGGGTCATTGTCCACTTTTTTAAACTTTTCATTT
GCTCATAATTTTCAGATTTACAGAAAGAGTTGCAAAATAGTTCCTATATGCTTTTCACCCAG
ATTCCACTAGTGTTAGTATTTACATGACCATGGTATATTTATCAAAAATAAGAAAAATAA
CATAGGTACAAGTACTATTAACTGAAGTGCAGAGTTTATTCAGACTTAAGCAGTTCTTCC
ACTAATGCCCTTTTCTGTTCAGCAACCAATCAAGGATATTGTCAACCGGGGTCTTGC
TCCCAGAGCTCCCAAGATGGTGGTGCTGCTCCGAGATGGGGCAAGCCTCATGTTCT
CTGACCTGGGGTCTTGGCCTCATGGATTACAAGGAATGGAATTTTGGGCTATGTGGTGA
GTGTTATAGCTCAATTAGAAGCCGTGGGTCACGGAAGAGAACTGTGGAAACCCAGTGACTA
GTGTTCAGCTCGATTAGGACCAACCTGGGACTTAGCTGTGCAGGAACAATGGCAAGCCTT
TAGCCCAACTTGGGAGCAGCAATGGGCACCTCGCTGGATCAGGAGCACAGCGGACACCCCT
GCCAGATCCGGAGGGATGGAAGTCAGCGGGGGTCTGCGAAGGCCGCAACACAGCAGCGGT
GGACAGCGAGCGAAAAGCTCAGCTCAAGCCGTAAACAACACGGACCAAGAGAGTGCAGT
TGCAAGATTTAATAGAAATGAACAACAGCTCCCATGCAAGGGAGGGGACCCAAAGTGGG
TAGCTGTTGCCGGCTCCAATGCCCTGGGTTTCTATCCCGATCATTTGTCCCTCCCGCTGTGC
TCTTAGGCAATAGATGATTGGCTATTTCCTTACCTACTGTGTTTTTGCCTAATTAGCATTTT

AGTTAGTTCTCTTTACTACCTGATTGGTCGGATGTGAGCTAAGTGCAAGCCCCCTTGTTTA
AATGTGGATGCAGTCAACTTCTCAGCTAAGCTTAGGATTCTTAGTCGGCCTAGGAAATC
CAGCTAGTCCTGTCTCAGTATCACAGTACATTACTTGTCTATGTCTTAGCCCTTGGCCA
GTTTGTAAAGTTTCTTGCTCCTACCTTGCTTTTCATGACCTTGCTATGAAGATGAAGC
ACGCTCTCTTGACTTCAGTCCTCCACCTTGGCTTCCCAAAGTGCTTGGGATTACAGGCA
TGAGCAACCACACCCAGCCTATAGTTTAACTTTAAACAAGAACGATAATGATCCCTTCC
TGAAACTAACCCCTCCTTGTTTGGGGACAAAACCGTTTTTGTAAACTAGTGGAAAGTC
CACAAAGTTTAGAATTTTGTAGGAACCTGAATTTCTGTAAGATGCAGGCATAAACTCTAAC
CAGTTTTTGCCTTTATAATCTGCCTTTTGTAACTGCTTACTACTCCAGAGTCACATAGC
CAGTGGTCACAAGATTTATAACATCTCCAATTGTCACTATAGATAATGTCACTATTGTAA
AACCTAAACTAGTGTGTAGTTATTTTTCAGACTTTACATTCTGATGGACCAACTGGTG
CCACTGGGACCAGTAACCCATAGCAAAAATGGACTCAGTTGATCCTGTGACCCCTTCA
AGAACCGATTCAACACAAGAAGACAGTTTCAACAGCCCTATGATTTTCATCCCTGACCCAG
CCAATCACCATTCCTCCTAGCCTCCTGATGGGCAGGAGGCTAGAGAGGTTCAAG
TTTTTTCATCCTTGAAAGACTGTAGCCTCCAATTCTCAGGAAGGCAGATTTGAGAATTAT
CTCTTATACTTGCCTTGGCCACCCTGCAATTA

SEQ.ID.NO. 409 HG1014958N0 PLT00014330H18.contig
ACCATCTTGGAACTCTAAGAAACAAGACCCCGCGTAACATTTGGTGGCCCGTACGGCG
CTTCTCCAAAGCGATACATCCTGGGAAGGACCCTACCCAGTCAATTTTATCTACCCCAACC
ATGGTTAAAGTGGCTGGAGTCTTGGATACATCACGCTCAAGTCAAACCCCTGGATA
CTGCCAAAGGAGCCCCGAAATCCAGGAGACAAATGCTGGGTATTCTCTGTGAACCTCTAGAG

GATCTGCACCTGCTCTTCAAGCGACAACCATGAGGAAAGTAACTAGAAATCGTGGATCCCC
 ATGGCCCTCCTTTGTCAACATTTTCTTTTACTGTCTCTTACCCCTTTCACTCTCACT
 TCACCTCCTCCATGCTGTACTACCACTAGCTCCTCTTACCAAGAGGTTCTATGGAGA
 ATGTGGCTTCCCAGAAATATTGATGTCCCATCGTATAGGGGTTTTTCTAAAGGAGACCCC
 ACTTTCACCAACCACAACCATATACCCCTGCACCTTCAGGCCATACATTTCAATCCCTGTA
 TCTTTAACCTCCTTGTTAAGTATGTCTCTTCCAGAATCGAAGTTGTAAACTACAAATGG
 TTCTTCAAATGGAGCCCCAGATGCAGTCCATGACTAAGATCTACTGCAGACCCCTGGACC
 AGCCTGCTAGCCCATGCTCTGATGTTAATGACATCAAAAGGCACCCCTCCTGAGGAAATCT
 CAACTGCACGACCCCTACTACGCCCCCAATTTCAGCAGGAAGCAGTTAAGAGCGGTCGTTGG
 CCAACCTCCCCAACATCACTTGAGTTTTCCTGTTGAGAGTGGGACTGAGAGGCAGTACT
 AGCTGGAATTCCTAGGCTGACTAAGAATTCCTAAGCCTAGCTGGGGAAGGGGACCGCACC
 TACCTTTAAACACGGGGCTTGTAACTCACCTCACACCCCAACCAATTAGGTAGTAAAGAGG
 ACTCACTAAAATACCAATTAGGCTAAAAGCAGGAGGTAAGAAATAGTCAAATCATATAT
 TGCCTGAGAGCGTGGGGAGGGACAATATTCGGGATATAAACCCTAGGCATTCGAGACGG
 GAGTGGCAACCCCTTTGGGTCCCCTCCCATTGTATGGGAGCTCTGTTTTCACCTCTACT
 AAATCTTGCAGCTGCACACTCTTCTGGTCCGTGTTTGTACGGCTCAAGCTGAGCTTTCG
 CTGCGCGTCCACCACTGCCAATTTGCAGACCCGCCACTGACTTCCACCCCTCCGGATCCAG
 CAGGGTGTCCGCTGTGCTCCTGATCCAGCAGGAGCGCCCTTTGCCGCTCTGGATCCAGCTA
 GAGGCTTGTCATTGTTCCACATGGCTAAGTGCTCGGGTTCGTCTTAATTGAGCTGAACA
 CTAGTCACTGGGTTCACAGTTCTCTTCCATGACCCACGGCTTCTAATAGAGCTATAGCA
 CTCACCACAGGGCCCAAGGTTCCATTCTTGGAAATCCGTGAGGCCAAGAACCCACAGGTCA
 GAGAACAAAGTCTTGCCGCCATCTTGGGAGCAGCCCAACCCCATCTTTGGGAGTGGCCTGC

CACCATCTTGGAGCTCTAAGAACAAAAGACCTGCTGGTAACAGCTGGATAAGCCCCAACTG
 CTTTGTAGCTGCCCTGTGTGGCAAAACAACCTAACACTCCTTAGGCCTCAATTTTAAATT
 TATAAAATGAAAGCAGTAAATATCACATCATATAAGTTGTTGAGTATTTATGAGAAATTAA
 AAGGAAATAATCGCCATAAAAGGCCCTGATACTGGCAATAATATATTGACAAATATGAGTT
 TCTCTTTTCCCTTTATTCTACTTTACGGCAACTCCTACCTCCTATTCTCTTTTTC
 TACTAAGCAGACATCTGGGTTTGTGTTGTTTTCATCAGGGAGAGGTTCTAAACGTTGT
 ATCCTTTAACCCCTCAAAGCATCATCTATTCTCTTAGAGGGTAAAGGGCAGAGATGTTATC
 CCTGTTGGTTTCTAACCTTGGAGAAAGTCAAGAAATGTCACGAGCTCTCTATAGATCCTCT
 TATCAGCTTCTCTTACACAATCCTTTTGGTTCTGTGTGGTGAAAACTGGATTGGGGGAA
 AATCAGATCAGTAGAAACCAAAATGTTGAAGCAACTGGTGAATAGAGAAAATACAGATTGCG
 TATCCCTACAAATGGAATGTTATTCAGCAATAGAAAGAAATGAAGTATGCACCCCATGCATG
 GAATGATGCACGGATGTTTACTACAGCATAGAAGGACCTTGAAACACAGGGCTAAGTGA
 GAGGAGCCGGACAGGAAGGATCATGTATTGTATGGTCTGCTCATGTATGATCTGCTCACA
 TATTGTGTGATCTGCTTACATGAAATGCCCCAAGGTAGGCCAAATCCACAGTCAGGAAATAG
 ATTGGTGTGGCCAGGGCAGGGACTTTTGGGGGAAATAGGGAGTGACTGTTAACAGGG
 ACCTGCTTTCTTCTGGGGTAATGAAAATGGTCAAACATTGATTGTAGTGAGGGTTGTGCG
 GCTCTGTGAATATGCAAAAACCCATTGAATTGGACACTTTTAAATGGGTGCAAAATTATAT
 TCCAA

SEQ.ID.NO. 410 HG1014960N0 PLT00014330I11.contig
 GAAAGGTAAACTGGTGGGGCGATGGGGTTCAGATATTCTCTGAGACCGGGAACGTCACT
 GCTCTCCCCGGGAAGCTGCTTTAGGCACACAGTGTCTCCTGCCTGACTCGGGGAAGACTGCTCT

CCTGGCCTCCTGGAAAGTCATATGGGAAACAGCAGATACCCAGTTTTCAGCTCCATTAACTGAAAGACTCTTTTCCTGAATGTTGTCTCTTGCAATCATTTGCACAAAATCAGCTTACCATAAAGAGGATTTACTGAAATCCTGAAACTTAGGAGTCAGCAAAGTTTTCATTTGTGTTGTCA AATGTCAAGAAAAGCTGAGGAGTTACAAGAAACTCAGTGATCCCTGACAGCTCTCTGCAA TTGCAGCAAAATCTGAGCAAGACTGACAGAAACCACCTCGGGGACAAACATCTGACTTCCA CAGCAGTTAAAATCATTTGAATACAGAAACCAAATGTCTTCATGTGAAAAACTCCATCAAG ACCCAAGAAAACTTCTGACTCAATTAGTAGTTTAGGGTACATCATGTGGATGCTTTTGAAA CCCCATTCTTTGAAGCTCTTCCCCACTTAACATATATCTAAAAACCATACATACATTCAAT AGATATAATCTTAGTTACCAGTTTTAAAGGAAATAAAAGCATATAGAACACAATAGTTAC ATCTCCAAAAATAAAGAAAGCTTGT AAGATGTGTAAAAGTTAAGAATTCATTTTCACTAACA ATTATTCTCAAACTTAATGAGGATACTGAGGACATGAAAATTACAGAGCTAGTATCATG ATGCTATAAAATACTGATATATTTAATATCCCAACAGGGACTGATTAGACTTATATATTT AAAACATATTTAAGTTGCTTAGAAAATTTTAAGTCAATAAAATTCACATACAAATAATATAC ATTATAAAATATGAATTTAAAGCTTTTGGAAATTTACTGATCAGACATTTTCCAATAGCTCTT TGAAATGTTCCCATACATATATTTTGTGTTGTGCTTGATGGCAAAGTTTAAAAAATGTTGA GTAAAAATATTAGAACTTGTGAAACCCCTGTTTGTATACACCCAAACACACTTACATATCT ATCTATATCTATCTATAGAAATATATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATTT TGTGTGTATGTGTTTCTGTCTGTCTTAGTTTCATTTGAGCTGCTATAACACAAATGCCAT AGACTGGGTGGGTTATTAAACAATAGAAAATCTATTTTCTCAGTTCTGGAGGCTGGGAAGTCC ATGATCAAGGCCTCAGATTATGTGTCTGGTGAGGTCTTGTTCCTGATCCTAGGTGATGA TTAATTGCTGGGCTTTCACATGGTGGATGTACAAAAAGGCTCCCTTGGGCCTCTCTTAGA AAGGCATTAATCTCATTCATCAGGCTCTAGCTTCTGGACCTAATCACCTTGGAGAGGCC

CATGTCCTAACACCATCACCTTGGATTAGGATTGCAACGTATGAATTTCACTGGGACACA
AACATTTAAGACCATACAGCCTCCTTCTTCCCTTTTCTACCCCCAACACCAATGCCCATC
CTTCCAAATTTGTAGTATTTCACTCACAGTTTATTTATTTTCTTTCATCCATTT
TCCTTAGACTGAATACATCTTTATGGGTCTCTCCTTTATGCTAATAAA

SEQ.ID.NO. 411 HG101496IN0 PLT00014330I12.contig

AATTGATCTGATTACTCAGGACTTTTCTTAGCCAGGACTGGGGGCAGAGAGGGGCTGCAG
TTGCTTTTCAATTCTATTTAATTTCTGTTTTCCAAAGATGAATATCAAGGGAGCAGCTTTG
ATTAAACAAGCCTTTAAAAAGGGGGAGAAATGCCAGGTTTTACAGAGAGCCTTGCTGAG
TCGCTGGCAAATGGCCACATGACACAATAATTTTAGTCTTTGGATAATTATCTCCTCTCTA
AGTACAAAAGCTGCTATTCCCATTAACACAGCTCTTAGGTTTCAGAGAAAAAGAAATTAGC
CTGGTGTTAAAAAGAAAAAAAATCAAGATTCAGGAAATCATCCTCTTCATAGAAATT
AAAAGAAAATTCTCATGTGCCCTAGTTCTACAGTTCTCCACATCCTTTTTTATCTGCTGCT
GTTTTTCTTTTTCTTTCTTTCTGCTAAAGTTGTTGAGACTGTGGGCCCCAACTG
GAGAAAGATTTTAAAGCCAAGTCTAGCATTTATTACATTCAAATATTTTAAAAAATGTTTT
ACAGGATGGTTTTTTGTGTGCTTTTTTTGTTTTAAGATTCTCTAAGGTGGGAATGTTGG
TCTTACTAAGATGGGAAATGAACAAGAGAGAAAAACCCCAATATTTAACATCGAACAAATAT
TTATTGAGCATCTATTAATGTGCTAGTCACTAAGGATATAGCAGTGAATAAAATGAAGCCC
TTTCCCTCATGGATCCCCTTCTAATGGGACAGCAGTAGTTAACATTTATTGAGGGCTTGCC
ATGTACCAGGCATTGTTCTAGCACTTTATATGATAACCCGTGGGGCATGCCCTAGTATTA
TTACTCTTTGCAGGCACTCACATGGCTACAAAGTGATGAACCAAGGTTTCACATCCAGGTA
TTCTGGCTCCAGAGTCAACCAACTTAATTTATTACCCTAATTTGTGACTCTTGTATTCCA

GATGAGTCTAAGAGTCCTTCTTCTCAGTTCCTCGTCATATCCTGTATGTATCCCCCTTTGGAGCCTCCCCAAACCTTTGTTCCTCGGTTGACCATAAAGATTTTGGAGTGTAGGGCCTGTCTGCATTCTCTGTATCCCAAGGCCTTACAGTGCTTGGGACCAACCAGAAATTGGAGATCCA

GCTCCTAGAGCATTTCAAAACACATGTGGAAGGACGCCCTGATTCAGGTACATACAGTAAATCTCCTCCCCAACCCATGAGTCTCTAAAACCCCTGTCA TTCTATCATATGTAGCTTAAGGTGCTGATTTATGGATAATGCGAGGATTTCTCTGCAGGGAAGGCTCTGCTCAAAAGACAGAGGACTGCTCTGTGTGGGCAACTTCAGGCCTAGGCAATTGCTGTCAGGCCGTGAGCTGGAGTGTTAATGGGGCTTGGGATTTAAACCCACACCTGGGAGTACTGCCAAGGTGCTTGCCACAATCTGTGGGCTGGGAGAAAGGTGCTCTGTGGCATCACCCAAATAATTTCCCTTAAAAATGAGCAGAAAACCTAAATGATCAGCAAGAAAGATTTAGTAAAATAAACCCCTGTGGTAGATGCTGTGCCATCGTTAACTAACCGGCTATAAAATAAGACTTAAGAATAGGGGTAAATGCTCAAGATATAACACAATGAACAGATATTATAATTTTGATGATAACATCATTTGCTATTTTGGAGGGTATACGTTTGCATAGAAAAGAACAGAGGCTAATCACCAAAATGAGATTTAATAGGTGGCATAATTCATTTCTTTTATACTTTTCTGCTTCCTCAGTTTGTATAAAATATGGTCAATACAACTAGAAAACATTATTAAATAATAACAAAAGGAAACTTTAAAAATGGCTTAAAAACGTTATATTGGCCTCAAAGGAGACCITTTAATCTAATTTTCTGTGCACACTGTCTTCTGTGCTGGAGCTCAGCAGCGATCATAAACACAGCAAATCGCTGGGTTGGAAACCTTGGCAGGTGTTTCTTGTATTCTAGACTTCAGGGGCTGTGGCTGTGGTTGCAAAACCTCCACTGATGATCTGATCCAGAGATCTAAGGCCTTGAGGATAGAACATCCCTCATCCTAGCAGGAAATAATGAACCTTCTAACCTTAAAGCCCTTTGGAGAGAAATGAATATAAACATGTGAGTCTAGCGTTGCTGCGATCCTTCATATTTACGGCACGCATTCTCATTTAATCCTCATAAACATGAAGGAGCAATTGGTCAATCAGCCCTAAGTTATAGGCAAGAAAGGCACCTTGT

CCCAGATCTTGTGGTGAGGACCAGAGCTTAAGCTAGACGCAGTGGCTCATGCCTGTAATC
 CCAGCACTCAGGGAGGCAGAGGCAGGTGGATCACTTGAGCTCAGGAGTTCAAGCCCAAGCC
 TGGGCAACATGGTGAACCCCTGTCTCTACAAAAAATAATAAAAAATTAGCCAGGTG
 TGGTGGCACACACCTGTAGTCCCAGCCACTCGTGAGGATGAGATGCAAGTATCACTTGAG
 CCCGGGAGGCGGAGGCTGCAGTGAGCCATGATAGCACCACTGCACTGCAGCCTGGGTGAC
 AG

SEQ.ID.NO. 412 HG1014963N0 PLT00014330I13.contig

CCGCCGCCGTGCTGCTCCATGGCCGCGCCCGCCCGCCCGCGCGTCCCGTCGCC
 GCCGCCCGCGTCCCCGGTGAGCGCTGGAGTGGGTGCGGCCCGTGTGCTGCCCTCCCGC
 GCCCTCTCCGAATCCCGGACTCCTAGTGAAGGAAGGGGAGGGCGCGGGCGGGAAGGG
 AGGTGGATCGGAGCGGCAGCGCGGCCAGCCGAACGGGTACCGACCGTCACCCCCGGATC
 TTCTCTGCAGGGGCGCGCGGCCCAAGCTGCTCGCCCGAGGGCCCAAGACCGCTTCCC
 AGACCAGAGCCGGGGAGGCGCAGCCAGTAGCGAGCCCTGAGCGACCTCTCTCCCCACTC
 GCCCCCTGCCCCAGTGAATTTCTGTCAAGTTGAAGGAGCGGGGCTCTTGCCCGGGGAAA
 AGGAAAACAGTTCGTGGCCTGGGATCTGGGACCCCTTGCCGCCCCCAAGCGCCCTCCCGC
 TCCCCGGCGCCCTTGCGGCCCAAGTGGCTCCTGCGCGCGGAGTCCGCCGGAAGACTGAG
 CGGCCCGGGGATCGCCGCGGGCTCCGCGCGTATTCCTGGGCTCAACACTGCCACTATCT
 GACACAACAGGAGCAGTGTGCGGGGGTACTGCGCGCCCAAGACGTGAGCTGTGATAGGCT
 CTCGAAACTGGCGCATTTTTTTTTTAACCTGGTAATTATATTTCACTTTCAGGGTTTTT
 TTAAATTATTATTATTTATTGTTGTTAAGCGACTTGATGCATTCAGATCTCA
 TGTGCATTTTCTCGTTATTCAAGAGAAATTGGACACTTGATGAAGGGAAGGACACGGAGAG

ATGACAAACGCAGACTGGCATGCAGTTGTTTTTCAGGGTGGGAAGAACAAAGTGGAATGTC
 TTCAGTTCTTACAAAATATGGTGGCATTTGTAACTCTGAGTTTTTTGTATTTGGCCTACA
 AAGGGGTCTTTTCAGGGAAACACGAATGGAACCCCTTCAGAGACATGTCTACATTGTAAA
 ACAATATAGGATAGATACGGATTTCGGGAAGGTCTGCAAAATGTTCCGCTCAAAAGAACTTG
 GGGTAAAGTTGACCGAGGACTCGGATTCTGAGAGGACTCCTATCACGTTCTTTCATGAGAG
 TTTTAGGCTATTTGGACTGAGGAATGATTTGGAGGTACGTGGGGGCAATCCCTCACTGA
 GGCCATCAGTGGCTGCCCTCCCTGCAATTGGCTGCTGGTTTGAGTTCCCTTCACAGAACTTCT
 TCGTCAAGTTTTTCATAGGAACGCTGCAGTGAAAGTGGAGTGAAAGAGACTGGGGATTGAA
 GGGTGAACCCCTTCCAGCCCCCAATCCCATCCCGATTCTCTGTCTGGATGCAGATTTCAGTCT
 GACTCAGAACGGATAGTTCAACCTTTTACTGATCCAAGTGTCTGTCTCTTTTTTTCAG
 AGGAATGTTCTGACTGGCTACAAACTTTTCTTCTGTCACTGCAGGACCTCCCGTGGAGCA
 ATTTTCGATAGGTCTTCTACTATGCTTTTTTACC CGGTCATTGTTATTGCTGAG

SEQ.ID.NO. 413 HG1014965N0 PLT00014330J10.contig
 CAATAGGCCAGTGACAGAAACACTGGTACAAAGCCCCTCGCTCAGTTTGCTGGAGAAA
 AGCAGCCGAAAGAGTCAAGCATGAGTGTGAGGAAGAAGAAAGAAAGGTGCCCTTAGAA
 AGAAGACAAAGGGCAACTCCAGTCTAAGAGTCTCAACCCCGCATAGAAAAATACTTTT
 GGGGCAACAAAATGGGTTGGGAGAAAGAAAGTATTTGACAGGAGTGGAAGCGGAGGGGCAG
 CTCTTCTGTCTCCTAGGCACATCTCCCTACCCCTGGAGAACCTGTCTGCTTGTGGCCTGG
 GAAGAGACTTCAAAATGGCGCCACCAACACCTGAATCAATAATTACCTAAGCGCGGTCC
 TACACAGGACAAAGAAAGGCAGCGTTTCCGAGAGTCCCTGGGATAAGGGGTTGCTTCTCC
 CAGCTGCGCATCTACATTTTAAATGTTGTGATTTACTCCGAGGGAAGCTTTCACACAAGTA

ACTAAATCCTAGGATCCCAGAGCTGCCTTCTTCTCTGGTCACCAGCTAAAGTAGGGGTC
AAATAAATAATTCCAGATTTTCTTAAATAAAAGGCCATATTTTCAGGTGGGAAAGAA
AGGTTGGTGGCTCCGTGTGACACAAAACACTACAACCTTTGGAAATGCCTGTGCCTTTGGGGA
AGGAACCAAAACCAAACTCGCAGAATACCTGCTGACTGAAGACAGAAATTCGCCAGT
GCTCTCTTCTCATCAGGGAACATCTTCCAATGGCCTTTCTCCAATTTCTCCTCTCCAA
GCCCATGGCAAGGGCCAAATGGGAAAAGCGAAAACCCCTAAAAGCGGAGCCTATCATTTA
TCCCAAGCCCCAAAGAACAAAGGCTATGTGGGTACAGTGGATTTGCTCCAGAAAGCCCCGG
CCTCTGCTCCTCGTTTCTGAAAACCATCCTTCTGGCAAGGAAAGAAAGTCTCAGCGCT
TTGGATGAAGCTCTCTGTTAGGCCACCACCTGCTTCTCAGGCCCTTCGACTGCTGGTGCAGG
AAGATTCCAGGCCACTGGGCTTACCGGCCCTCAATGCTCACTTGTCCAGCCACAGCTT
GACAAACTGGTCACTTCCACCCACCTGCTGGAGCTGACTAGGGCAAGTAAGGCTGATGG
GTGACCCCTGATCGGCCCTCACTGGAGTACAGGGTTTCCCTTCTCCAATCTCCTCAAAACAG
GTGAACAAAAGTTCTCAGGCCAAGTTGCCCTCCAGCACCTCCCCAGTCTCAGACAAAGTCCAT
GGACCATCTCTCCTGCTCCATGCAGGAGGATTCTCCAACCCACAGTGCCAGTTCTACAGGT
GTTATCACCCAGCTAGGAAGGATTTTGAGAAAGACAGCTCTGTCTATGGGCTGTGGGTT
CACATCTCCACTCTCCACCCAGACAGCCCCAGAAAGGCTGTCTCCCTTTGTGCGGGCTG
AGGTCTGAAGAGTGTGAGGCCCGGACACACACCGGTTCCCAATGGTCAGGGCTGCTTG
GGTGATCAGGCCCTGGTGATGCGTTTCTGGGCTCCTGCCCTTCTCTGCGGGGTGACTCCCCG
GCTGTGCTCACAGCACACGCCTAGAGATGCCCTGAGAACACTGAGGAGCACACAAGGGG
CAGCACATGCCTCTGGAAACCTTTAGTTCAGCCTTGTGTGGAGGGAGACACCTTAAATGAAT
GAGACCCCTTTGGGAAGAGTCACTGCCACGGTGTAGCCAGGCTGACAGAAAGCCGTCCTGTA
GAATTGGGTGAAGGGTCCATTTACACAGCAAACCTGGAATGTTTCCTAAAGACTGAAGAGGA

AAACGCTGATTTATTAGCTGGTGTGAGGCCAACTCTACTTCTATTTCCAAATTTCGTCT
 TTCCTCTCTCATTTTCTTCATCTTCTGTACTTAGGACTTGGGGCTTGGGGATGGGGACGC
 AGTGGCTGGAGGCTAGACTTTCCTCTTTGAAAGGTCTGACTTCAGCTCCCGTGACTATTG
 CTAACCCGGCAGCCACACTCGGGCCGCTGCCTTGCTGTTGTACCCCCGATACGCCCTG
 GCCTCTGGGTAGGCAGACGCTTCACTCTGCAGCATGGATGGGCTCAACCCCCAGAAAGCC
 AAACAGCCCTCCCGAAGTCCTTTCTTAATTCGTAATTCTATCTACGTTTCTTTATCATTT
 TATAAAGCAGATGTTTTGGATGATATGTTTCCCTTTTGTTCATTTTGTACTGGGTTGAATA
 TCATCCCCCATAAATTCATGTCCACCCAAACCTCAGAACGTGACCTTAATTTGGAATAG
 GGTCATGTAGACATGATTAGCTCAGGTGAGGTCATATTAGATAGAGTGAGCCCTCATTTCC
 AATGACTGGCGTCCTCACCAGAAAGGCCACGCAAAAGACTCAGAGACAGGAAGGAGGCCAC
 GTGGCCACACAGGCAGAGACTGGTGTGATGCAGCCACCAGCCAAGAACGGCGGCAGCAGC
 AGGAGCTGGAAGAGGCAAGGAAGGGTCTGTCCCCGGAGCCTGCAGAGGAAGTGTGGCCCT
 GCTGACACTTTTGACTTTGGACTTCTGACCTCCAGGGATGTGAGAGGGTAAATTCCCTGTTG
 TTTCAAGACACCCAGCTTGTGGTACCATGTTACAGCAGCCCCAGGAGGCTAATAATATCC
 TTTCAATTTTGTGTTATGTTTAGACTTATGAATATAAACAAATACAATAAAAAA
 ATA

SEQ.ID.NO. 414 HG1014967N0 PLT00014330J14.contig
 ACCCTGACAGGGGACTTTTCCACCTGGCAGTGCCAAATATCCGGCTCCTACGCATTCTCCA
 TGGAGGATCACCAAGTCCAGTGCCCTTTACACGGTGAAAACGTGAATACAGCCATCAGT
 AGAAAAAGAACCCCTCTGAAAACAGGAATTTCCAAGACAGTGCATTGAGGAAGTCCAAGG
 AGGAAAGGATGGGGAATAGAAAGACAGCAGCGAAACAGGACTGGGGACTGTACAGGAGCCA

GAGAAAGGGCTGATTCTAAGAGGAAAGGAGACACCAAAAATGTAAAAGGGTGCAAAAATAC
 TGGAGAGAGGAGGCCCTTTATGATGGCCACTGGGTGGGGCAATTGGCAGCCACCAAGGCCT
 CTGAGAGAGCAGAAAATGTGGTTCCACCAAAACACTCTTCTGGAACCTGCTCTTTTGTG
 TAAGGTGAGCTGAGCACTTAGAAGGATCTTCCTAAATAATCCCCCTTATGAAGAATTCTA
 ATAAACCCCTTCTCTGTCTCATCTTCATTAAACCACTAATTTTAAATTTTAAATTTTAAAT
 TTTTATTGTGTAAGAGAAACACATAACATGAAGTTGCCCTCATACTAGATGTTTGC AAC
 ATGGCATAGTATTGCTAATTACATAACATGGGGCTGTACAGCAGCCCTCCAGGGCTCTTCC
 ATCTCACGTGGCTGAAACTCTCCACGTTCCCTCCCCCAACCCCTGGCAACCACTATTCTAG
 TCTCTGTGTCTATAAGTTTGACTGTTTTAGAGACCTCATATAGGTAGACTCATACAGTAT
 TTGTCAATTTTGTGACTGGCTTATTTCACTTAGTATAATGTCCCTCAAGGTTAATCTGTATT
 GTAGCATGTGACAGGGTGCCCTTTTAAAAGGTGGAATGATAATCCATTATTGTGTATATA
 CACTTTCTTTATCCATTCACTTCCCAGTGGCCATTTAGATTGCCCTCCACCCTCTTGGCCAT
 TGCAATATAATGCTGCAATGAACACACAGTGTAGATACTTCTTGGAGATCCTGTTTTTCATTTT
 TTTTTCTGATGAATACTTAGAAGTGGGATTGCTGGAGCATATGGCAGTTCTGTTTTTTAAG
 GCTTTGAGGAACCTCCATGCTGTTTTCCATAGTGGGTGAACCATTTTACATTCCACAGCA
 GGTGCACCATTTTATATTGCATAGTGGGTGCACCATTTTAGTTCCTGCCAACAGTGCACA
 GTGGCTTCAATATCTCCTCCTCCTCACCAACACTTGCTACCTTCTGTGTTCTTGATAATGG
 CCATCTTAGTAGGTGTGAGGCTTATCTTACACCATACACAAAATAAACTCAAAATGGAT
 TAAAGGCGTAGCATAAGACCTGAAAGTATAAAACTTCTAGAGAAAACAGGAGAAAAACT
 CTGTTGCAATTGGTCTTGGTAGTGACTTCTTGGATAAGTCCTAAAGCACAGGCAGTGAAA
 GCAAAAAGAGACAAGCAGGATTGCATCAAAACCCAAAAGCCTCTGTGCAGCAAAAGGAGGCA
 ACGAGCAGAGTGAAAAGGTGACCTATGGACTTGGAGAAAGTGCTCTGCAAAACCATGTGATA

TGGTTTGGCTGTGCCCAACCCAAATCTCATCTTGAATTCACCTGTTGTGGGAGGGAT
 CTGTTGGGAGGTAAATTGAATCATGGGGTGGGTCTTTCTGTGCTATTCTTGTGATAGTGA
 ATGAGTCTCATGAGATCTGATGGTTTATAAAAATGGGAGTTTCCCTGCACAAAGCTCTCT
 TCTCTTGTCTGTGCCATGTGAGATGTGCCCTTTCACCTTCCACCATGATTGTGAGGCCTC
 CCCAGCCCATGTGGAACGTGAAGTCC

SEQ.ID.NO. 415 HG1014969N0 PLT00014330J15.contig

GCAATACTGTTTGGAGGTTTCAGACGAGGTGTCCATTTATATTGCCCGACAATGATGAAAGT
 CATCTACGGAGGCCCTCTCCAGTTTCATCTGTACAGGGCTTTATGAACCTTTCTAACCAA
 TGATGAACCAGAAATGCTGTGACGTCAGGAGAGAGAAAAATCAAAATAACCCATCCAAAGG
 GACCGTAGAGAAAAGTGGCTCCTGTACAGGACATCGCTCACAGTGTCAATCAGCAACAAG
 ACTGTGCAACAGCAGACTCAAGCTGTGTGTTCTTGTACTGATTCTTACACACAGTGCT
 CACAGCCTCGGCAGCACAGAACACAGCCGACTGAGCTTTGGAGGCATCAACACGCTGGA
 AGAAAACTCAACCAATGAGGAGTAACGGAAGGACGAGTGTCAACACAGCAGCAGCTGGCC
 CGCCGTGAAAAATGGCAACTGCTGTCTCATGTAAACAGAACTGGGTGCTTTTACCCCTCGA
 ATTACTTATTGCAAGGCCTTTAGGGTAAAAATTTAAACAGATGGGCCTGAATCCAAACAAG
 GACACAACCAACAGCTTTTATTGACTAAAAGGCTGGAAGTGACTTTAAATTTCTCACAC
 CATTTTATACACTGTGTTTAAATGTTTGGAGGTTTATTGCTTTCGTTTGGTTTGGGT
 TTATTGTTTGTATTATTTTGCACTTGTTAATACAGGATTTATTTGGGGGATGGTTTC
 TCAGAGGTAAACTAAGTCTTTTCACTGTCTCTATCTCTCTATATAATTTCTAGTCATTGIG
 TGTGTTTCATCAGATAGTTCTGTCTTTATGTCCTGTCAGCTTCTATTAGAGGAATGATTGC
 TATGACCTCATGGTATAGCAAAAAACAACAACAAAAAGAAATAAAAAAAGACA

AAAAAAGAAAAACAATAAAAAATAAAAAATCCCTAAGTCTCCCTTCTACCC
 ACAGAACCAACAACACCTTCCCGCCTTTCCTTTCCTCGCCCTCTTCTCGTCCCTAA
 GCAAAACAATCCGCTTGCTGTGTGTAAACCACAGTGAATGGGTGTGCACGCTTG
 TGGGCCTCTGAGCCCCTGTTCACAAACCAGAAACAGAGCGGAGCCAAGGGGCCCTGACA
 AGAGTTCCTTTTAGCTGAACAAACAAGTGTCTCCATAATAGGTGGAATCAGACAGTTA
 ACACATTTTATGTTGAAAAACAATAAAAGGAAAAAATTAATAAACTATCATGAAC
 GTATTGCTCCAGTTCCTCCATCCCCAAGTGGCCCCAGCCCTTCTTGCTGCTCCAGTTGGACA
 GGAGCAGCTATCTAGAAATCAGGATGCGGGAGTGAGGAAGTTTCTTTTGACAATGAAG
 GTGGGCTTTCATTTGTGATTTTGTCTGTTCAGTAATATAGGAGCACATTTTGGCCATT
 GTAATTACAGGGAACAACAAGGATTGCGGACACATATCTGGACTTCTTTCCCTCCCT
 TATTGTTGGAGAGACACTAGAAATGCTCAACACCTGCAATATACAGAATATACACA
 ATTTTATTCCAGTATTTCCCTAACATATGGTTTAAATTTATCCAGGTATACAGTGTATG
 CAATTCTGCATTATCACAGAGGAACAACCTCTTTTAAATAAATAAGGTCAGCCATT
 TTTATTAAACGTGCAAAAACITTTATCACTCTAACATGCTCTAGGTAGTTGAGGAAAAAGAGG
 TCTGATCACTGTTTGTATTTTATTTTCTTTGTGGGAACAATTCACCTGCTGAGTGTACAT
 GAATTTGCTTCTATAAAAGGCTTTTATGAGTTTACAGTAGAATCAGTGGAAAGGAAGAGT
 TAATAAGGGCTGTTTAAATAAAACAACAACAACAACAACAATAATAATAAAAAAAA
 TTTTACATTCTCTAATCTCTAACTACACTTGGGAAGTGCACTTCAGATAAGTTTGCA
 GTGTGACTGAGAGATGAAGGAAATCCATAGAAAAGGTCCTCTAGTGAACAAAAATTTAGT
 TATTAACTTTATAGCTATGAAATTTCCCGGGCATTTGTTTTTGTTCACACAGACTTTAA
 CCTCTGCATCATCTTAACCCCTGCGACATGCGTACAGTATGCATATTTTGTTTTGAAAAA
 AAATGTTTCCGTTCAGTCTGTTAAGAAATATTCAAAAATAATAAAGGTATTGCTTAATAAA

ATTGCTAGAATTGTTTAGCAGTACATGCACAATATTTTACTAGATTCTTTGTTTAAATAG
TGTTTGTGAGACTGAAAATCTTAAATGGTCTG

SEQ.ID.NO. 416 HG1014970N0 PLT00014330J21.contig
CCACTAGAACAAAAATGCGGACTCCCCTAAATAGAAAAACAGGCAAGAAAAACAATCAC
AGAGTGAATTTAACATAATAGAAACAGTAAAAATAATATAAAATAATATGACAGGCTTAA
GACTAAATATATCAGTAATATAATATATAATGAGCTGAAAAATCACCTAATAAAAGGAA
AATATTTTCAGATTGACTTTCAAGCAAAACCAATTCATTGCTCTGTAGAGAGACATACC
TAAACAAAAGGATTCAGAAAGGATAAACTAGTAGGATGACTTAAAGGCATACAAAAGAT
GCTCTTAGCAGTGTTAATTATAAGACCCCAAATAACTACCAAAATGCTCTTCAGAAAGTAGA
ATGAATAACAAAGTGTGGTGATTTCATACAGTGGAAATACTACGTATATAATGAGAGTAT
GAATATGAATTTTACAAAAACAACATGGATGAATCTCATAGTAAGGAGTGAAAGAGGCATG
AAACAAGAGTGCAATCCTATATGATTCCATTTATATGAAATTTTAAAGAGGCAGGACCAA
TCTATGGTGATAGTAATCCGTATCATGGTCACCCCTTATGGAAAGGGTAATGACTGGAAAGGG
AAGCCAAAGGGCTTTCAGAGCACTAGCGATGCTCTATTTTATCTGGGTGCTGGTTGC
ATGATTTTGTAACTTTGTGAAAAATTTCTTTAAGCTGGATACTTAAAGAGTTGTGTTCCTTT
CTGTATCTACGTTACATTTTAATAAAAGGTTTACATTTAAAAATAATAAACTATTTACAAG
CAGAAAGATGAGGCTTGGGGTAAAGTCCCATGGAAATAGCTGGACTTAGAGTCATCAGAGAG
AATGAATTTTAAATCCAAATTTCTATGAACACTGCCTCTAGGCTCTAAAGGCCATTTTCTT
TTCCTTTATGACACTTCAAGGTGTGCTCCCTAGATTGTGTTTTTGCCATCTCCAAATTAATAA
TAGAGTCTAGAAAAACAGGGTAGAGAAATGCAGAAAGTACAGCCGAATACACAAAAGGCATTT
GGCATAGAGTAAACCATGGTATTTTCAAAATTTGGTGGGATGGGAGAGAGGAATGGATTTTTC

AATAACTGGTTTTGAGACAACCTGGGTAGCCATTTGGGAAAAAATTAAATTGGATCCCTGT
CTTCCTCCTTATATGGAAAGTAAATTTTATATTGATTAAAAATATACATGTACAAAAATGAAA
CCATAAAAAATATTAAAAAGAAAATATAGAGGTATATATTTTAAACCATGGAATATAATGT
AAGCCTGACTGAAAAATTCAGGAGCTATGTAAAGAAAAGATTGACAAATTTTAAAAGTTCTA
CATGGTTAAAAATAATATCATAAAAACCCAAAAGATGAACTGGGAAAAAATATTTTGCAACACA
TGAGAAAGGGCTAATTTCCCTTTATAAAGTGTTCACACAGTCAACAAAAAGTCAAAAAAAGA
TCAGTAATCCAAAGAAAAAATGAGTAAAGAACATGAACCTGTTGTTTAAAAACAGACTAACT
AAATAACTTACTAACACCCGCAAGACTTATTGCAACATGGAAACCTATGAACATAGGAA
ATGAATGATGCCAGGTCTTATTTCATAAATAGAGGAATACAAATTAAATTGCAAGTACAGTA
TGTAAGTCATTTTTCACATATTAGATAAAAAACATTTGCTCATACATTGCATTTCTGAGTAC
AGAGTAGCAAGAGTCTCCTATATTGTTGAATGGGCAGTATTAGAAAAATAATTTAAAAATGT
ATATACCCCTTTGACTTACTGGTTTCCTTCTACCCTTTATTCTGTAGATACACCCACATC
TACATTCATTACAGCAATTGTTTTATGGCAGCAAAAAAATTGGAAACAATGGCAGGTACCAACC
AAAAAAGAGACTGGTTGAATAAACAAAAAGCATAGCTATGCAGTGGAATGCTATGCAGCT
GGTAGAAAGAAAGGGCAGATTGTATATGTACTAATATGAAAAAATCTCCAAGACACACTGTT
AAATGAAAAAAAATGTGGCACACCGTGCCACGTTTTTAAAAAGGGAAAGGGGCATATACACAGA
TGCTTGTTTACACAGTTAGTATTTCTTGAAGGATATATAAGAAAAATGGTTTCCTCAGAGG
TATAGACCTTTGGGGAAGGGTAGGGATGAGAGGGAAACCTTACTTGTCAATTGTTTACTGTT
TTGTATTGAGAAATGTTTCCCCCAGGTGAATGACTGTTATACCTATTTAGGAAAAACAGACG
AAACCAATAAAACCAAACTATTTCTATACAGTGATTATTAGCAGAGATGCCATTGGCATT
TTAGATGCATTTTTGGGACTGTTCCACACATTTAGGACATTACATTCCTTTCCCTTTGCAC
TAAATGCCAGTAGCACTTCCCTAGACACTGACAAACAAAAAGTGCCCTGTAGATTTCTG

AAAAGGCCAGGGTGTGGATTGGGGTGGTGGTGTGTTGTTGAGTCTAAAAGAATGTAACTTC
 TCTTCAGGGTTTCCAGTTTCCAGGTTATGTTGGCCTGGCAGGAGGGTCTTTGACCGACA
 CGGATGGTTATGGACCTTTACTTAGTGAAAAAACTCAAATGAACCTTCCTTTCTAGAACTCA
 TGGTTTTCATTCAGTGACACTGGACTTTGCCCTGGAGCAGTGACTGTCAGGGAGCTGTA
 TCTAAGTGGGAGTTGCACCTGTGTGGGCAAAAGTGTGGCACTCAGATTTTCGTAAACAATG
 GGCAAGATTTCAAGGAAGCAGGATTTTGGTTCCCTTACTTTAAAATGCATGGAGAAATGGCT
 GGGTTTGGTGGCCCATGCCTGTATCCCAGCACTTTGGGAGGCCGAGGTGGGCAGATCAC
 CTGAGGCCAGGAGTTCAAGACCAGCCTGGCCAAACATGGCAAAACTCCATCTCTACTAAAA
 ATACAAAGCCTAGCATGTGTGGAACATGCCTGTAGTCCAGCTACTTGGGAGGGTGGCGCA
 CGAGAAATCACTTGAGCCCCAGGAGGCAGGGGTTGCAGCAAGCCGAGATCACACCACTGCAC
 TCCAGCCTGGGC

SEQ.ID.NO. 417 HG1014972N0 PLT00014330K01.contig
 ATACCTTCACGATAGTTGCTTCCAGAGAAAAGGCATGCAAGTCTAATTTAATGGTAAGAA
 TCACTTACTCTGTACAGTGGTGTGCTGAAGGAATTATCCTGAAGGCTGAAACAGAAAC
 AAGAAAAACATCAGCCAAAGTCTGGCCGGGGAGCCGTCTCCCTGATTGCTCTCCGGGAGTGA
 GCACATACACAAACATGGCACTTGCTTTCTTCCACATGTCTCCTACCAATTTACAGAGC
 TCACCCCAACCCCCACACTCCCCACCCACCCCGAGGATGGAGGCTCGCAGAGATTTCAT
 CCATCGTTCAGAAAAAGTCAGGAGAGAAACAAGAGGGGAGGGGGATATTGCTGCTGTTT
 GGTGTTTTATTTGTGTGTGTTTTTTTGTGCTAGGAAAAAGAAATAGGCTCTAAAGCA
 GGGCTGGCTGCATCACACCTGATCTGTCTAGAGCGCTCTTATTTCAGTCCCAGAGAGGAA
 GCGCCAGCCCCGGACTACAGAGCTGTGAATGAACCCACGCGAGGCGAGTGGAGAATTTA

TTCCAGGGGTTAGCTGGTGATTAGGACACCGTGCAAGGGGTGGGCTTGCTATGCCTAGA
 GCAGCCCCGCTTGAGCTCAGCTCCCTAGCTAGCACCTTTACAAAGGATGGAGAGACGGAAT
 CCGCAAAACGCTATCCCTTTGAATAAAGGCAGCCCTTGTTAGGGAAAGGAAAGACTGTCAGT
 CCCCCCACTGGCTTCTTTCCCTTGGGGATAATTGTTCTGCCCTCTTTGTGATTGTTCAATTT
 CTCCTGCCATGTTGACCTGGATATGGGATGTAGCATCTATTAAAGATTCTTATTTGAAATT
 TGAAATTATTTTGAATTTTGAATAAAGATTCCCATCCCCCTTTTAACTCTTCTACT
 CCCAGAAAGCAGCTAGGATAGTGGTTTTTAGTCCAAACCTATGTAAAGTTAATGAGTTAACA
 TAACAACAAAAAAGCTATAATTTTCTAATTGGCAATTAAACCTTTACCAAAAACCTTGG
 AAATAATGGGGAAGAGCATGACCTTTCAATGTCAAACCTGAGACATCGGTTAACAAAGTCCA
 CCCAGGGGCCAGTTAATTCTCCAGTTGTGTTTGGCTTTGCTCTTCACTATTGATTAGATT
 CCAGCCACAGTGAAAGTTCTGTGTCACTTGCACATTTTATTTCAGCCTGTTTCAGAGAAAAA
 GATAATCGCAGCGTTTATGATTTTATTGCAATATAGGAAGATAACCAATTTTGCATCTAA
 CCAATAATCATTTTGCATCTAACCAAACTTTCTTCTAAATACCTGTAAATCCACAGAT
 CCTCAAAATGCTTTTGGAAACCAGCTTCATCATCTTACAGTTTTCCTCCCATTAATAAGTGAA
 TAGATCTCTGCTGCCTGCCTATGTTTGTATGAGAGTTGTATTTTAACTTTTAAATTTAT
 ACCT

SEQ.ID.NO. 418 HG1014973N0 PLT00014330K08.contig
 AATTCCTCAAAATGGTTGTCTCACCTCACTTCATGAGCTTTCCTTTTGTGTGTCATC
 TTTGTGATCGGTGAAATGAATTCGGGTGTTTGCTCATGAAATCGAGCATTTTTTTCATA
 TTAATTGGCCTTTTATAATTTTGTGTGAATTGTCTGTATTACATTTTACCTTTTACATT
 CAAATACTATACAGAAATATTATCCTGTACAGCTATTTTATTTTGTGCTTACCTTTTAT

TTATTTAGAGACAAGGCCTCCCTCTGTCAACCCGGGCTGGAATGCAATGATGCAACCATAG
 CTCACCTGCAAGCCTGGAACCTCTGGGCTCAAGTGATCCTCCAACCCCTGGCCTCTGCTAGGA
 CCACACATATGTGCCACCAATGTCTTGCTAGTTTTTAAATTTTTTTGTAGGAACAGGATCT
 CACTACATTTGCCAGGCTGGTCTCCAAGCCTTGAGCAATCCTCCTGCCTCAGCCTCCCCAG
 AGTGTGGGATTACAGGTGTGAGCCAAACACACCTGGCCTTTTTTTGTCTTACCTTTTCAATT
 CTGTTTTGCTGTTTTTTTGACATATAGGGTTTTTAAATTTTGTGTCAGTTGGATTACATTGC
 CTTTTGTACTTTTTTCAATTGCTTCTATAAAGCTAAATCTTCTTTTCCCATTCCAAATTC
 AGAAATAATTTCTACTAATTTAGTTTTTCTTCCTTTTTTTTTTTCCCTTTTATATTTAGCTC
 TTTAATCTCCACAGAAATTAATCTGTGTGAATCTTATGAAGTAGGAGTTAATTTTACTTTTT
 TCTAGAATAATTAATCTGTTCTTGTGTTTTATTCAGAAATCCCAGTATAGCTGTGACATGAAC
 CAAACAAAGTGATTTTTTCTTATTTTAGTGCTTTGATTTTCCCAATTAGGAAAAATGAGGAA
 GCTGTGAGTACCAGTGGAAAGGAACGAGCCAGGAAGAGGGACTGGAAACCATCTCAGAAAGCC
 ATGCCCTCAGGCTGGAACCTTGCCCTGTCTCCTCGCAGATGAGGTAGGAGGCTCAGCACG
 GGCTGGTGGAGCAGCCCAACAGCAGCAGTGAGGGTCAGTGGGCCCTGTGAGTTAAGAAAGAA
 GGTGGCACAGAGCGAGGGTCTCTGGATCCTGACTGTTTGACTTTTCCATGATTAGGATAA
 GTAGCCAGGGCTTGGCTAGTTGGAGAGTTACTCGAAACCTCAGGTGACAGTTGTAAAGGCAG
 CACATAGTGAAAAAGAGTCCTAGCCTGGGAAAGTCCAAAACCTTAGGTCTGGTTTCAGTT
 CACTCACCTATCTCTGTGACATAGCAAAGTCTGTTAATTTCTCTAATTTTCTGATTAGTA
 GTATAGTTGCAGGAATGAAATAAAAAATGGTCCTGATTACTCAGAGATTCAATTACTGATT
 ACTCTTTTTGTGACCTGAATAAAGAGTACAACATCTCTCAAAAGGTAAACAATATGATGTT
 TCAGGAAATTAGAGAAAGTTAAGAGACTTTGCTGTATTTACTATACAGATAGTTGACTGAC
 AGCTGAGCCCGGACCCAAAGATTCAATTTATATAAAATGAAGGAGTTGAGAAAGTACATCTCTG

AGGTTTATCCTAAATAAACTGAAGTCTAGGCAAGTGGCTGAGCTAGGTTTGCCATGAATC
AATCTGCTGAGTATATTCTTGGTAAC TAGTTCATCTTTCCCTTAATTCAGTCAACAAATA
TGATATCAAGTTTAGTAAGATGATGACAGTTGT CATGTAACTCCATCTGG AAGATTGT
ATGCTTCTCAGCTTCCCTAAC TTTTCTTAATCTACCTTAAGCTTCCCTTCTGAAAGGTTT
CTGTCTTTCCCTAACAGTCTCTTCTTTCCGCCAGCTACCTCTTGTGTTTAAGAAATTAAAGG
AAGATGAGCTAGGAATCAAGAAATACGCAGAACTGATTCTCTGTCACTGGCGCCACCCTGC
CACTTGCCCGAGAGCTGTGGTATAAATTTTTTTAATAGGCTGAAATAAATAATGAATAGC
ATTTTGGTATTAAAGCATTA AATTGATAAAGGCTATGAGATACACCTGCTCTCAACCATGT
TAATTTTTTTATTATGGTATTAATAAATTAATTACTTACTAACATATTAAACAAGAATTGCA
TTGAGAAACAAAGCATCCACAGGCCAAATCTTCTGAAATTCAAAATGTTTATATTAAATGCG
TTGTATTCTAGAAAGTAGAATTGTCTTAAAGTAGCTTTGTAAATATAAAGTCATCTATCAG
CCCATACACCTATTAGAAATGTTTAAACCTTTATTTCCCACTTTTTTGTCTTAGACTGA
GTGTACTTGGTCTGGTATTTGTGGCATAACTTACGGAACACATAAGAATGATACAGATAC
TATTTAATGATGACCTAATACAAAGCTTAGATAGCTAAGGTGAAAGCTTCTATGGCCTTAA
CATTTTCCTCTTGAAGAATGTATTTTCTGTAAATAAATA

SEQ.ID.NO. 419 HG1014975N0 PLT00014330K09.contig
CTGCCTGGACCTCACCATAAGTCTCAGGTATTTCTTCATAGCAGCATGAAAATGGACTAA
TACAGCAAATTGGTACCAGGAATGGGGCGCTACTATAAGGATACCAAACATGGAAGCA
ACTTTGGA ACTGGGTAA TAGGCAGAGGTTGAACAGTTTGAAGGCTCAGAAAGACAG
GAAAATGGGGAAAGTTTGGAACTTCCTAGAAACTTGG AAGATTTCAGAAAGACAGGGAGATG
TGGGAAAGTTTGGAACTTCCCTAGACACTTGTGTAATTGCTTTAACCAAATGCTGATAG

TAATAAGAACAATGAAGTCCAGCCTGAGGTGATCTCAGATGAAGATGAGGAACTTGTGG
GAACTAGAGTAAAGGTCACTCTTGCTATGCTTTAGCAAGAGACTGTCAAGCTGTTTGCCCC
TACCCTAGAAAATCTATGGAACCTTGAACCTTGAGAGATGATTTAGGGTATCTGGCAGAAAGA
AATTTCTAAGAAAGCAAAAGCATTCAAAGTGGTGACAGAAATAAAAGTTTAGAAAATTTGCA
GCCTGATGATGCAGTAGGAAAAGAAAAGCCCCATTTTCTGGGAAGAAATTCAAGTTGGCTGC
AGAAATTTGCGTCAGTAATGAGGAGCCAAATGTTAATCACCAAGACAAATGGGGAAAATGT
CTCCAGGGCATGTCAGAGACCTTTCACAGTAGCCCCCTCCCATCACAGGCCTGGAAAGACTAG
GAGGGAATAATAGTTTCTGCGCCAGTTCTAGGGGCCCCCTGCTGTGTGCAGCCTCTGGA
GTTGTGCCCTGCACCCAGCCCACTCCAGCCATGGCTAAAAGGGGCCAAGGTACAGTTCT
GGCCATTGCTTCCGAGGGTGCAAGCTCCAAGCCTTGGCAGCTTCCACATGGTGTGGTTC
TGTGAGTATGCAGAAGACAAAGAACTGAGGTTTCAGGAAACTTTGGCTGGATGTATGAAAAT
GCCTGGA TGTT CAGACAGAAAGTTTACTGCAGGGATGGAGCCCTTGTGGGAGAACCTCTGC
TAGGGCAGAGCAGAAGGGAATGTGGGGTTGAAGTCCCGCCACAGAGTCCCCACTGTTC
ACTGACCAGTGGAGCAAAGAAAAGAGGGCTATAGTCCTCCATACCCCAGAAATGGTAGATC
CACCAACAGCTTGCACTATGCATGCACCTGGAAAAGCCATAGACACTCAACGCCAGCCCA
TGAAAGCAGCTAGGAGGGGAGCTGTACCCCTGCAAAGCCACACAGGGGTGGAGATGCCCAA
AGCCATGGGAGCCACCTGTTGCATTAGCGTGACCTGGATGTGAGACATGGAGTCCAAGA
AGATCATTTTGGAACTTTAAGGTTTAAATGACTGCCCTATTGGATTTCAGACTTGCAATGGG
GCCTGTAGCCCCCTTGTTCGGTCAGTTTCTGCCCATTTGAAAACAGGTGTGTTCCTCCAAT
GTCTGTACCCCTGTTGTATCTAGGAAGTAACTAACTAGCTGTGATTTTATAGGCTCATA
GGCAGAGGGGACTTACCCTGTCTCAGATGAGATTTTGGACTTGGTTTTTTTGGGTTAATGC
TGGAATGAGCTAAGACTTTGAGGGACTGTTGGAAAGGCATGATTGTGTTTGGAAATGAGA

GGATATGAGATTTGGAGGGGCCAGGAGCAGAAAGATACAGTTTGGCTGTGTCCCCACCA
 AAATCTCCTCTTGAAATTGTAGCTCCCATAAATCCCCACATACTTTGGGAGGACCCAGTGG
 GTGGTAAATTGAATCATGGAAGCAGTTACCCCCATGCTGTTCTCGTGATAGTGAAGTGTG
 TGTATGCCCTGAAATGACTTTTCATATTCCTTTTGTGCTCCTAGAACCTAATTTTGGTTTTTG
 GAAAATTCAAAAGAAATGGAGAAATCAAAATCAAGAGATAGAAAACGTCATGCTATTCTGATAT
 ATAGAACTCAAAACAATTTTGTGTTGA

SEQ.ID.NO. 420 HG1014977N0 PLT00014330K15.contig

GGGGCCGGCCGGGTAGCGTTTCTTTTAGTGCCGTAGGCAGCTCTGGCTCGGAGAGCCT
 TTTGCTAGCCCCACGGGACCTCTGTGCACGGATGGACCCGCCGACCTGGCGGGAAGC
 GGCTGGCAGGCGGCGCCCGGCGCATCAGCAGAGACAGGACGGGCGCGACGCGCGG
 GCCCCTGAGGCGTGCCTGCCCAACCGGGCCGGCGGCACCATGATGCCGGGCGAGACC
 CATTCGGCGGCCCGGACGGCGCGGACCTGTGCGGATGTACAGGCTGCGCCTCTCTG
 CAGCAGAAATTTGAAATGAATATGTTGAAGCATTAAATTACCTTGAAACAAAATTATCAAT
 ACAGATAATTTGTTAACAGAAATATCAGAAGAAATGTGATGATATCCTTTTACTGGCTTTT
 TCTGTTGAGTTTGGAACTAACTTATGTTTAAATTCCTGAAAGAGATGTGAGAAATTGTTCC
 TCATAGGCCCTTCTGGAAGATAGTACTTAACAGGCTTCAGTAGGGAGCATGTCTCACAAGC
 ACAGTAAACAGACAGTCCTTTTATTCTTAATCCATGTTTAAATGATGAGTTTAAATAATGA
 GATTACTTTTCTTGTGATTGTGATTTTCTTAACGTCCTCATAAGCTTTTATGTTGGTGGG
 TTTCTCTTAAATTTGTACTGAGGATGTGGCATGTTTTTTTAGGTTTACGAATAATGTGAAAC
 TGAATGTGATACCTTGCTGAATGTTATTTTGAATCTGCAAGAAATACAGTATGGGAAATT
 CAAAATGTTGGCAATTGACTTTCATTAAATTAATTTTGTGATTTACAGGGAATGTGTGT

GTATATATATCTCCAAAATCATAGAAATAGCGTTGCTGTATATTTCATGAACGCTGTGG
GCAGTCAGAAAGACAATTCTTGGCCCTTACAGTCAACCACCTTGTTCGGTAGCCATTTAATGAG
CTCGGTTGTGGAGTTAGTGTCTGCTCAGCTATTTTTTTAGCAGTGGGAAAAAGATGTGTA
TAATTGATCCTTCCATGTCTTGCTGTGGTAATAACCTTGATTAAATCTCTCCTGTAAAAA
CTTAAAGTTTCTCTCATAAAAGTAAACATACGTATACTCCCTTTAGGGAGTCCCTGTAAATT
TAGCTTCATGTTTCGATGTCAGGACAGATTACAGCTGTTTCAGCTTTGGCAATAAAGTTGAGGC
CTGCTTCAATATGATTTATGTCTATACCTGAACCTTACATGTCTCATCATCAGTTATTTATTG
TGTTGGTACCATCTGTACTAGTTGATTCTCACGCTGCTATAAAGGACTACCCGAGACTGG
GTAATGTATAAAGGAAAGAGGTTTAATTGACTCAGAGTTCTGTCATGGCTGGGAGGCCCCA
GGAAACTTATAATCATGGCAGAAAGGGAAGCAACACATCCTTCACATGGTGGCAGGAGA
GAGAAAGTCAGAGCGAAGCGGGGAAAGCCCCCTTATAAAACCATCAGATCTCGTGGGAAC
TCACTCATTATCAGGAGAACAGCACTGGGGGACTGCCCCCGTGATCTGTATCACCTCCCAC
AAGGTTCCACCCAGCACCTGGGAATCACAACTTGTTTACAAATTCAAGATGAGATTTTG
GGTGGGACACAGCCAGACCATATCATCATTTTCTAATTTCTGTGAAAAATCGTAAAAATTAA
GAGATATCCCTTTAGATTTTATTTTACTATCTGAATTTTACATTCAAAAAATAAAGGTCAT
TGGAAATCTTCTTATGTGCCTCATCTGTAAACCATCAGTCCATTTAACTATACGTGCA
TATTATGGCAGAAATGATATCTTCAATTGTTAGTGAAAGGTGTATGGAAACAATTATCAC
AACTTTTAGAAGGCTAAATAGGGAAATAGATGGGCAGTTCAGTATTAGGAGAGTATGTG
TTCATTGTATCAATCTTGAAAAAATATGCACACTCTTTGAGTTAACTTATACATGTAAC
CTAGCAGAGAAATAATTGTTTTACTTTTCTAAAGGTTGGCCAGAAAAATTTTAAAACAAA
AATTATCAGAAATCCTTTATTAAACGGATAACACTTTTGATTATTCTCGGATTTTCAGAAAGTT
CATTCTGTAATATGATAATTTGGCTTCATTGTGCTCAGCTACAGATAGATACATACCTTTAAAAA

AAATAATGTTTCACATTTTGCTCCAAAGTATGCATTTGTGTACCTCAAATATGCAGTTGTTTT
 GCTCTGCAACTTTGAAGTGCCGATGTCACAGATTTAACATGGCTTTTAAAAATATAAATTTG
 ATTCCTTTTGTCTACTAAAGTAGCAGTGCAATTTCCCAATTTTTCAGATGGGAAAAGTTA
 TTTCTTCCATAAAGTGCACATATGATAAATTCAGTATTTTCAGCTCTTTTGTGTAAATTCCTGA
 CCACCTTTATATCACCTCTTTCTAAATGGCAAGAAATGTATTTTATTTTCATCAGGTATTT
 TTTGAACTTTTCCTATTTTGCAAAGATTAAATTTAGGATGAATAAAATATAGTTCTTGTCT
 TCAGAAACCTTCAGAAACCTTACATTTCTAAATGGAGGATGCACACGTTTTTATGAACGTACA
 TGTCTATATTTGTGATAAAAAATACAGGAATTCCTAGGAGATGAAACATAAAGGAATTTGTT
 TTTTGCTATCTTAGAATAATTTTACTAGCAGTATGAGATGGTAAATAAGGTTATCAAAATT
 ACATAAATTTGTAAAAATGAGAGTTTTGTAGATTTTATGATCTGTTTTCTACAAAACATG
 CTTATATATGGTTTTTGTAAGTACATAAACTTTTGTGTGCTG

SEQ.ID.NO. 421 HG1014979N0 PLT00014330K24.contig
 TAAGTGCGCACTCTGGGCCCTTTGTGTATACGGTATGCACTGAGATATAAAGAAGAAAAA
 GCCACAGCTCCTTCCTTCGTGATGCTCATAGTCTGGTAGAGGAAGAATTCAGGTAAGCAA
 AAAAGCCAGTCATATGATCATCTCTTTCTGCAATTTTGCTCAGGATCAACTTCTACACGAC
 TGATCTATAGCTTATGGCCCTGTCTCTTTCCCTTTCTTGCTAAGTCATAGACATACCAG
 TTGAATTTATACAAATTCATCTGTGGGAAAAGTTTCACAAGGCTGATTTTTCATATATCT
 TCTCTGTATAAATTTATGTTAATAAATTCATTTAAAGTATTTTAAATCCTAAATATATG
 TCAGTAAAGGTGCTAAATGTTAAACAAATAATTACTTTTCAGATTTTCTTTCTTCTCTCT
 TTTATTCCTATTCTATTTTCCCCCTTGGTTAATTTATTTTCTCAAAATTCCTCCCTTC
 TTTTCTAAACCTTATATTTCCCCCAAAAGAATAGTACCAGATAGTCTCATATCAATTTGAT

CTTTTITGGCCTGGTTCCTTATATTGACTTTCCTGTCCAAGTCTTTTCTACTTGTCTTAAG
ACTTCATCTCAATTTTATTATTGGCCTTTCAAGATTGTGTAGACAAACAGTTTTCAGTG
AAAGAGCTTAATTTAGTTAAGCACTAAGTTTTTATCTGGGCTAGAAAATTTTGGATCTTAC
CTTTACTCTTTTCTCACTATAACCCCAAAACAACATTCAATTCATCAGGAAAATCTGTCAAT
ATTTCCCTCTTGAAGCAATTTACATCTGTTTCCAAAGCTTTATTCCCATATTTAGTGTC
CAGTCCATGCCCTTATCTTTTCTCAGCTGATAATGGGATAACCTCCTAAATAATTTCCCAT
TTCTTCTGACATTAATCTTAGTCITTTTGTTTTTCAAAATTATGATTCATGTTCCCTGTAGGT
TATAAACTCAATTTACTTGTTCATGATCATAAATTTTGTAGTGATGTAAATAAGAACATAT
ACACAAATATTGGAGTACATGGCACATAATAAGAGATTTTCATGAAAATTGTTTTAATTATTT
ACATATGTATGCTAAGTGTTAACTTACCACCTGTGTAATGGGAGATTATAATTGTGTGGTT
GTCCCTTGGAGGGCATCTATCTAGCATCTGAACCTTTCCCCCATGATTAGAAATAATTCTGT
ATTATATGAGACTTGTAGGATGTAGGGCCTGCTTCCCAATTATAGAAGCTAAAAAATAGCC
CAGGTCTTGTCTTTCCCAAGCATCCCTTGCAGCTAGGTTGCTTACCTGTAAACCTAAGCTTG
CTTAATCAGATGCCTGGTTGTGAACTGGCACTTACAGAAAGCAGGGAAAATGAATACCA
AGGCTGCTGGTGTGGGGACACCAAGCAATTTGGCTTTTGAGGCAGCAGCGGCTGACAGCC
AGTGATAGCAACAGGGTCCCTCATCAGAAATGGTTCTGGTCATGGTTTGGCTATGATTCTGG
CTACCTCACTTCCTTTTGTCTCCTGCCTATTTTCTGAGCTGGCACTTTAAACTTCCTGGAG
ATTTTATGAGCTACTCTAGTTTCTTTCAACCAATTTCTTTTCTACTTGTCTATGGACTGAA
TTTTGTCACTCTCTACTCCTCCATTCAATATGTTGAAGTCCTAACCCCAAGTGACTAT
ATCTGGAGGTAGGATCTTTGGAAAGTAATTAAGGTTAATGAGGTCATAAGAGTGGGACCC
TAATCAGATAAGACCAAGACTTTTATAGGAAGAGGAAGAGAGAAACCTCTCTCTT
TCATGTGAGGATGCAGCAAGAAAGGCAGCTGTCTGTCAAGCCAGGGACCTCACCCAGAAACCA

AACCCTGTCAAGCCTTGATTGTGGACTTTCAGCCTTCAGAAATGTGAAAAA

SEQ.ID.NO. 422 HG1014980N0 PLT00014330L01.contig
 TTGGGAAAAGATGGGGGAAAATGTTTTTGGCTTAGATGAAGGAGCAAAAGAGGGTGATGAG
 AGCAGGAGGGGAAAAAAGGGTGGCCAGGGATAAGGGGAAAAGATGGTGGGAAAGAAACAGT
 GGGGAAGGGGTTTGGGTAATGGATGGGAAAAGGTGTTGAGCAGGAGAGTAGAGAAAGGC
 TTTGTGAAATGATGGTGGGAAAAAAATGGTGAAGAAAGTTTGGGGGCAGATGGGAAGAAAA
 AGGTGCGGAGGGGAAGGGGCCAAAGGTGCTCGGGAAGAAAGGTAGGGAAATAATGGT
 GGGGGACAAATGGTTTGTAGTAGATTTTTTAAATAAGATCATTTTGTATTTTGTTTTGAA
 TAGTCTGAGATCTTCATATATTTGTGCAATGAATCCCTTGCCTGATGCATAGTTTGCAAA
 TACTTGCTTCCATTCTCTTGGTGTGTTCTTTCATTTTAAAGATTTTAAATTTAAATTAAT
 TTTTTTTAGATGGAGTCTCATTTCTGTCACTCCAGCCTGGGGCAGAGCAAGACTTCGTC
 TCAAAAAAAACATTAAAAAAAATGTATTTTCATTTGTGACTCATCTTGTAACCTAGA
 GGATCGAAAGGACTTTGAATGATCCTCCTGACAAAGGAGGGAAGCAGTACCTTAATCAAC
 CCTGGGTCCAAATTAAGTACTTTTATAAACCCCTTGNCACATACTTCATATAATTTGTGG
 CCTAGATACTGACTTATAAACTAAATTTTTTATTTTCCCTTTTATACCCCTTACAAAAATCA
 TATGCAGAAAGTCCAATATCTAGTTGGTCAAAAGATGTGCTGATCTGGTTTCAGTGGGGA
 AAGTCCGGGTGAGGAAGGGATCTGCAATCTTGCCCTTATACTCTGTTATTTGTTGAGATGT
 TCATGAAAGGCCCTTTTATAGAAAATCCGTATGGATGCCGATTACTCTTTATATTGAGTTGG
 CAAACTGTGTGAAGGCCAGATAGTAAATGTTCTGGGTTTGTGAACCATATGGTTTCTA
 TTGTAACTATTTCATCTTGCCCTTGAGAGCACCAAGCAGCCGTAATTAATATGTAAATGA
 ATAAACAGAGTTGTGTTCCAGTAAAGCTTTCATTTATGGATACTGAAATTTAAAGTTCATA

TAATTTTTTTTTTTTTAAGAAAGGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGT
GGGCCCATCTCGGTTTACTGCAACCTCACCTCCCGGGTTCAACAGTTCTCCTGCCTCA
GCCTCCTGAGTAGCTGGGACTACAGGCGTGTGCCAGCACCCCGGCTAATTTTTTTTTT
TTTTTTTTTTAAGTAGTAGTGGGTTTACCGTGTGGCTAGGATGGTCTCGATCTCCTG
ACCTCGTGATCTGCCCCGCTTGGCCTCCCAAGATTACAGGTGTGAGCCACCAAGCCTGGC
TGAAAGTTCATATTATTTTATGTGTCATAACACATAAAATTATTTTATGTGTCATAACA
TCTTTTTAAAGATGTGCTTTTAGCTTATTATTGGGATAAAATGTTTTGAATTTGAATA
TAGATGCCCTGTTTAAAGATTTTTTTTCCACACAGTTTGTGCAATTAAGTACCAGAATA
CCCAAAACATTTTTCATTAGTCTTCATCAACAACACTCCCGTTTGTGTTCTCAGTCAAAATTAAA
AATAGCAATTTTGGCCGGGTGCGGTGGCTCACACCTGTATCCACGCTTTGGGAGGC
CGAGGCAATTTCCACTAAACTACAAAAATTAGCCGGGTGTGGTGCGGGCTCCTGTGG
TCCCAGCTGCCCTGGAGGTGAGGGAGGAGAAATGGTGTGAACCCGGGAGGCGGAGCTTGC
AGTGAGCTGAGGTGCGGCCACTGCACTCCAACCTGGGCAA

SEQ.ID.NO. 423 HG1015004N0 PLT00014330L24.contig
GTCGAGGCTGTCGGAAGGGAAGGGAGTCCCCCAAAGACTCCAGTACTCATTCCTTCAAGC
CTCCCACTTAGATTTCTCTCTCCATCTCTTTGGATGCCCTCGGACAAATTCTCTGCTTCT
CCGAGCCCAAGTTTTTGCAATTTGTAAACAGTTTCACTGAACCTGTGATTGAGTTGGGGCG
GGGGCGGGGTGATACAGACTTTGGCCGTGCACAGCGTCTAGCTCTGAGTGAGGGCCTAA
TACTGTGGGGGAAGGAAGAGGTGCGCCATCCTAGGCTGTTAAGGTGGA CTCACTCTGC
CCCTGCCGTGTCTCTCTGGGAAAGCCACTGGGCTTCTCAGAGCCCCCCCCCCCCCG
CCCCTTTTGCATATCTGCACAAAGAGGACGATGACACCTACTTGAGGGGGTTTTTCTCTGGG

GATCAGCTGAGAAACAGGAGACTGGCTTAGCACAGTGCCAGGCACCTGAAACTATTGAGG
CTGGCAGTGGCTGCCCTGCTTCCATCTGCCCCCTGTCGTTCACTCCAGTGGGCATTTCACA
GAGCATCTGCTCTGAACCAAGGTCCATCGAGAAACATGCAGGGTTGTGGCAGAGGCAGATGG
GACCTGCACTGTGACTTGGCCAACGTGGACTGTGCCGATGTCTCCTTCTTAGAGCCTCA
GTTTCCCTCAGGAAATACAGAGCAGCTGCCTCCGTGACTGCTGGATTAGAGGAGATGGT
AGTTGAGGTACACAAGACCTGAGTAGGGAAACTCACTCTGAACCTCCCCGAGGTTCTG
TCACTGGCCCTTTCTCCAAGGTTGTGGTGACAGGGCTTCAAAAACCCAGGGGAGAAT
GGTGAGGGCTCGGGTTCCAAAAGCAGAGTCAATGGACAGTTTCCCCCTGTGTAAATAGAG
GGGAACAGCAGTGTGTTCTCTCCGGGAGGATTTCAGAAAGGGACAACACTAGCACGGGGTC
TGGGTAGGGTGACGTGTGAACCAAAGCCTTCCCCTAGAGCATATGCCGGGACCTGACC
TCTGCGTTCCCTTTTCATTTCATTCACCTCATTCAATTCATTCAATTCAGTGTCAAGTGGT
CTGACACTGTTGGGGGAGAAAAGGAAATCTCAGCCTGAACCTGGAAACCTAGGGACAGCAG
GGAGAGTTAAGGGATCTTTGAGCAGCCACTGTCTTCCAAACAGCCATTCTGTTGGCCCTGTGT
GGGAGGAAAACATGCAATGATGGATCTTTGACCTCGAAATCCTCCCTCCTCAAACTGTGGTC
CACAGACCAGCCACATGGCCATCACAGGACAGACTCTTGTCCCCCAGGCCAGACCTACTGA
ATCAAGTGTGCAATTTGGGATGATATCCAAGGTCAAGCAAAATGCCTGGCACAGAGTAAATGC
AAGACAGGGTTGACTATCATTTGCTCTTTTCGAGTGGGATAGAGCTCCCAACTTTTGGCT
GGGCATGTGGGTGCTCACAGTGAAAGTTTGCAATTTCCCCCAGCTTCTTTGTAGCTAGATGT
GGCATTGAGAAAGTAAAGGGAAGTGTCTTTGTGGGCATCTGGGATAGTTGGTATGTCTGCT
TTACCCCTCTGTGCCCTTTCTCCATCTGGCTGCCAGGAATGTAGGCGTCACCATCCTGGA
CCATAAGGTGGAGGCCACCACAGAGCAGAGCAACAAGACAAAAGAGCCTAAGTTCCTGC
CGGGCACAGTGGCTCACACCTGTAAATTGCAGCACTT

SEQ.ID.NO. 424 HG1014981N0 PLT00014330M02.contig

TGAGAGCCCCGCACTCTCACCTCACAGACGGAGCAGACAGATCAAATGGAAAGGCTCCC
 CCATCCTGTCCTCTACACCACCTTGACAGCTGGGCTCAGCAACTGGGCTTTTAAATTCAG
 TCTAATTCAAGTCAGCAGCATAGGGCAGCTCCTGGGAAATTGGTTTACACATGCGGGACAA
 GCCCAGTAGCCCAGAGCTAACCCACTCACCATCCCTGACCACAGAGGAGCAGATAAGGAA
 GCAAGATTTGTCGGGAGCTATGACAAGTTCCACTGTGGATGGCAAGGAAATGAGGGGGGC
 AGCAACCTGTTAATGAGACTCAAGAGTGTCAAGACCCATTTTGTATCTGGATTTTCAA
 GGCCGAGTTCAAACATGAGGATCCTGGAAAGGTTGGCAGACGTCTGCGGAAAGCTACCTTC
 TCTGTTGTAAAAGTCCACAAAATGAGAAAGTGGCTGCAGTGTATAAACCTGTGGACAGCTCT
 TCTTGTACTGTACCCCACTCTCATAGCTCACCATCCTCAGGCACCTCTTATCTCTGCTC
 ACAGAGTACCCACCTGAAATGCCTAACCTGTCTAACTCCACCCAGTGAAATTTCTGTGAATC
 CCACAAGGTGTGCTTATATGCCCTCCTTTCCCTCTGCTCATTCCTTCCACAGACATTATTG
 TACATGCCCTTGGGTGCTCAACGGCAGTCTCTGGATTCCCCAAACAGTACAGTAGTTTCTT
 TTAGCCCTTCCCAATTGCACCTTGTCAGGGCAATTGCTTTTGGCAATTCAAATTTGTGTCTA
 GTACATAATGGAGACCTGGCTACAGAAATATTGCAGTACTGAAACAATCTCTGCTATCC
 ACGGGCTTGAAATTTACCCCTCTGCAAGTGTAACTGGAGGAAACACAGTAGTCACCAACT
 ACGGGCAGTTACTGAAATTCAAATTAGTTGAAATAAAATAAAACCAAAATAACAAATCCT
 TATTCTCAGACCACATTTTCAGGTGCTCAGTAGCTACATGTGGCTGTGACTGTTATATC
 ACCCGGTGCAGAGAGAAAGCATTTCTACCCCTTACAGAGCATTCATTAGACAGTGTGCT
 CGGGATTCTATATCCCCCTGAACCTCTGCTATTCCCTGTCACTGGGTCGGTTTCCTTTAA
 AACCCTAACCAATTATTTCACACACTTATTTGTGTTTATGGTGATTCCCCCGCAAGTAGACTG

TAAGCTCTGGGAGAGTAAGGACTATCTTTATCTCTTTTACCACCTTCCACCCCTTACACTTA
 GAAAGAGACCTGGCAGGTGACAAATAATATTGCTAAAGTGAAACAAGTGGGATGGTGAGTT
 TCTATCCACTAGAGAGTCATAGCTGTCTAAAGGGACACCATGCACCATTTTTCAGCCTT
 ACCTGGTCTCTCCAGATGCCCGCAGATAACCTTGTGGTGTGGCCTTCGCTCTTGCAGGGTC
 TCAGCCCCCTGGGACATCCTCCTACATCCACATCTTCA TGGCTGAATCCCACCCAGACCT
 CAAAAATGAACGAAAGGCCACCTTCAGGAAGCCTTCCTCAATCCAGTGCTTCCCTTTTG
 TCTCCTCATCTTCTTCCCCCATTAAGGGCTTAAGAAATGATGTCTCTCCAAATTTGACCACC
 TTGGCTTTTAAACTGCAACTCTCTGTGGAAGCTTCACACCTGTTGCCCTTATTGGTGGCTGCT
 TTATACACATTTGTTCA TGTTCCTTATCCCACTACTTGAGAGCAGGGTCTATGTCTAGCTT
 ACAATCATCTTCCCAGGCTCTGAGAAAAGCATTTAAAATACAGCAGAGGCTTAAACATA
 ATATTTGAATGAATATATGAGTAAATAAATGATCATGTTTTAAAGCTATGATTTTACTAT
 TTTTAAATGATCACTACCACCTAAAGAACATCTTCTTTATTTTAAATCCATGGCTTTT
 TATTTTATTTTAAAGATTTCATACCAACAATAACAGCATAAAGCTCTATAACTTATAA
 AATTGTGTAACCAAAATCTAAAATTGAGAAATTTATTATACACAAAAGTTAAAGAACCTA
 GGGCTGGAAGAAAGCAAATTTCTATACTGAAAATGTCTTTAGGAATGGGCACGTATTCAAT
 TTATTTTGCACAAAGGAAATTTCAITTAATAAGAAAATAATGTGAACCATTAGTGTTAAC
 TTAATTTGGTTTTTAGATGAATATTAAATCTTATATCTAGAAATATGTCTTGCATCAGCA
 TTTCAA

SEQ.ID.NO. 425 HG1014983N0 PLT00014330M08.contig
 TACAGGTTCAAGGAAAAGTTAGTGGAAGCCGTTATCCATGTCAACATTGAACCAGAGAA
 ATAGAAAGCAGAGGAAGCAGTAACCTCACTTCTATGAGGCCAGCATTACACAAATATGAAAC

TAAACAAAGACATTGAAGAAGAAACTATAGAACACTATCAGTCATGAACACAGATCCA
AAATTCCTCAACAAATAATCAGCAAATGGACTCCAACACTGTATAAAATAATTATATAC
CATGACCAAGTACAATTGTTC AAGGAGCATCCTTCATGATGATTAGTCTGTGATAATT
TTGTCAAGCAAATGACACCTGTTATCACCTCCAAGGAAGGAGTCCCACTTCAGCTGAAGC
AGTAGTGCCAAAGTTTTAACCCACTCCATCTTTCAGGGTGGCCAAAGAGATAGGACTCTA
TTTAGCCTGGCAAACCCCTCTGCAGGAGTCTCCCCAAAAGTACAGTGCCCCCAAACCATTTCT
GAGATGGGGCCAGCTGGGATTCAAAGAAAGAACCAAAATATTAAAGTTGATTAGTTTG
AAGCATTCATTAGGAGAACCTTACTTACAGAAIGCTGCAGCATATCCTCAGGCAGAAAGG
GAAAGAAAGGAATGTTTTACCTGAATATGACTTCAGTGACGAGGTACGGTAGGAAAGCTT
ATATAAGAGTTTAATGAAGTTGGCTCACGGAGAGGACTAGTTTCTTTTTTCTTTTGAT
TTTTTTAAATCAACTTTTATTTTAAGTTCCAGGGTACATGTGCAGGATGTGCACATTTGT
TACATAGGTAAGTGTGTGCCACGGTGTCTGTGTGCACGGAACAACCTATCACCTACGTAC
TAAGGCCAGCACCCATTAGCTGTTCTCCGTGATGCTCTCCCTCCACACCCCTCTGAC
AGGCCCCAGTATAGGTCGTTCCCCAGCATGTATCCATGTGTTCTCATTCAGCACCCCA
CTTATAGGTGAGAACATGCAGTGTTTGGTTTTCTGTGTTAGTTTGCTGAGGATAATGGCT
TCCAGCTCCATCCATGTCCCCTGCAAAGGACATGATCTCATTCCTTTTTTGTGGCTACACAG
TATTCCATGGTGATATGTACCAAGTTTTCCTTTATTCAGTTTGTTCATGATGGCCATTG
GGCTAAATGCAATGTCCTTGTCTATTGTGAATAGTACTGCAATGAACATACATGTGCATGTA
TCTTTATAATAGGAAGATTTATATTCCTTTGGGTATGTACCTATTAATGGATCGCTCGG
TCCAGTGGTATTTCTGCTTCTAGATCTTTGAGAAATCACCAACAATGTCTTCCACAATGGT
TGAACATAATTTACATTTCCACCAACGGGTGTAAGAGTGCTCCTTTTTTCTTCACAGCCTTGC
CAGCATCTTTTGTCTTGACTTTTTATTTTATTTATTTATTTATTTATTTATTTATTTTGCAC

TCCAGCCTGGGCAACACAGTGAGACCCCTGTCTCAAAAAATAATAATAATTTTAAAA
 AAACAACCTTTTGCTCCATGAAGAACCTCATTAAGAAGATAAAATAGACAAGCTACAGGGTG
 AAAGAAAAATATTTACAAACCCTATATGCAACAAATTACTAAGATGTAGACTATATAAAGG
 ATTCATAAAACCTTAATAGTAAGCAAAAAATCCAATTATGGAATGAACAAAAACACAAAGA
 AGTAACTTCTCAAAAAAGAAATATATAGATGACAAAAATAAGCTTATAAAAAACATGTTTAAACAT
 CGTTAGCCATTAGGAAAAACGCAAAATTAAAAATCACAATGAGATATTACTATATATTCTCG
 GAATGGCCTAAAAAGTAAAAATATCAAAATGCTGGCAAGATATGGACAAACTAGATCAC
 ACGTACATTGATAATGGGAATGTAAAAATTATACAGGCATTCTGGAAAAATACTTTGATAGT
 TTATTTAAAAAGTAAGTGTACAACTACTTTTATGCCCCATTATTGTACCTTTAGAAAAAT
 TATCTCACATAAATGAAAAATATGTTAATACGAAAAACATGTGCATGAATATTTATAGCAAC
 TTTATTCACAATATCTCCAAAAAAGACACAACTCTGATGTTCTTCAATGTTGAATGTTT
 AAATAAACTATGACATATGTCTAGCATGGAAATACTGCTCAGCAAGAAAAAGGAAATGCATT
 ATTGTCAATGCACAAACAATCTGGATGGATCTTAAAGAGAAATTATGCTGACTGTAAAAAGCCA
 ACCCAAAAATGTCACATAAGGTATAATTTCACTTTTATCGAATTATTGAATGTCAACAGT
 TATAAAAATGAAAAAAGATTAAATGATTGCCAGAAAAATTAAGGAAGGTTGGGGGCATGAG
 GGAAAGCAGGTGTTTATAAAAAGACAACATGAGAGAGACTATTTGTTGGAAATAGTCTG
 TATCTTCACAGAAATCAATGTAAATATCTTGGTTGTACATTTTACCGTAATGTTGGAAAGG
 TGTTACCAATTGAAAGAACTGGTAAAACCTGGATAAAATGAAATCTCTCTGTATTATTCT
 TACAAAATACATGTGAATCTACAATGATCTCAAAAATAAAAAATTTTAA

SEQ.ID.NO. 426 HG1014985N0 PLT00014330M15.contig

GGGACAGGAGGGTCGTGGCGACACTCGCGCCAGCTCCCGGACGGCTGGGCTCGGGCTG

GTCCGACCTCCGACCTCCACTAGATGCCTGCGAGAGCCCGCGCGTCCGGAGCCCCGC
 GCCCCTCAGAGGGAAGGGTCCGCGTCCCGCGCGGCCCTGAGGCTGACGGTCGGGAG
 GTGAAAGCCGGAGTGCGGGTCAGGACGCCCTCCTGGAGGGGCGCGCAGGGGGCCG
 CGGGAAGCCGCCGAGGGCGCGGAGCGGGAAGGAAGACAGTGGCCACCCCGCCTCCGCC
 TGGCACGGGCGAGGACCCGCGGAGGCGGTGCCGGGAGGTCTCAGTACTCAAAAA
 CAAGCAGAAAGGCATCCAGGGCCCCCTCCCGAACTCCCCACGGGCTGCCAGGTGCAGA
 GAAGAAGCGCCTGAGGAGGAAGTGCCACACAATGGGAGAGAACAAAGGGAGCGGAGCGA
 GACCGCGCCTGCGCCCCCCCCCGCTTGCGCACAGTCCCCCGCCCCCGCGGGCGCC
 GCCGCCCGCCCTCCCTCCCGCGCCTCCTTCTGACCCCGAGCGAGAGACCTGGTCGGGA
 GTCCGGCGCTAGATTGCAAGGGCCTGTGCTCGACTGTGCGTTAAAGAGGCTA
 GAGGAGCCTAGGTAGACGTGGGGCGCAAGCTGGAGCCGCGAGCTTGCGAGGGGGCCTTT
 ACGTGATGAGGCTCACCATGGGGTCTCCAGCGAGGTGTCTGCAGAGCTTCGATGGTGAGA
 GGAGAACTCGATTCTCGTTTCTGGAGTTGGGATGTAGACATAGAGAGATCTGTTACA
 TCTGAGATGAAACCCCTGCACAAATTCACGTTAGGGTTGGCCAAATTGTTGCTTGACCTTA
 AATGGTTAGTGGTTTAAAGCTTCCCCAAAATGCGCTGTGAGCTGAGCTTCCACTTTGG
 AATGATTGCAGAAATTCGGTGCTTAAGAGATGAGAAACTTCGTTCTACCAGTAATGGAAGA
 AAGCTTCACTTAATCACCCCTCCTTGCTTTCCCTCCCTTGACTATGTAGGACTCGA
 TAACACAACATTTGAAAGATTCTACTCTATGAAATAAGCACACAGATTGAAATTCAGTATTT
 GCCAAGGAATAT

SEQ.ID.NO. 427 HG1014986N0 PLT00014330M17.contig
 CCCGACGCGGCAAGATGGCGCGCGCGCGCTGCGGCCCGTGACCTGGACCACTGGGCT

GGGACCCACGACCCACACCGCCTGGAGGACGCGCAGCGCGCCTCGGCCTTGGAAG
 GGGAGTTTCCAGCGGAGCCCCGCCCTTCCCCGCCCGGAGGCGCCTCGGGCGGGGTC
 ACTGGCCGGAGGCCCGCCCTTCCGCCCGGAAAGCGCCGGTGCAACCCCGCAGGCCCTGCC
 GTGGGTCCACAGGCCGGATCCCCGGGTTCCTAGGCTGGGACTGCTTCTCAGAAACCACTCT
 GTCGTTTTTAAGCAGGGTCACACACTCTAGCTCACTGGGTCCATTTTAATTTCTATTAA
 CATTTTTTTTTTGCAAAATGATGTAGTAGGAGATCCAAGGTGTTTGTTAATGATTTA
 TTCACATCATTAGTCATTCACAAAACCTTGCTTGAGCACCTGTTATGTACCCAGCACTGTG
 CTGGAATGCTGAGGAGACAGGAGTGAAGTAAAAAGACATGTTCCGGCAGGAACAGGCA
 AGGAGAGCCTTGACTTGTAAGGGCTAACACAGTGCTACACAGGAGGACGCCCGCTCGGT
 GTCGGGGTCAGCAAAGCTTCCTTAAGGACCACTAGGTTGAGACAAGAAGGATGAGTATG
 GGATTAGCCAGGAAAGGGAAAAGGGCAGTGGGAACTGGAAAGAGAGCCTTCTAGGACT
 GGGAAAGAGCAAAGTGCCGGGTCAGGAGAGAGGACATAGCAGTGAGGAGGCAGTAGGTATG
 GATGAATGAGCAGAGGAAAGTGGGGAGAGATTAAAGTGTCTTTCAGAAACGGGTTGGAAACC
 AGTTTAGTGAGTTGGATGGCATTTACCTAAAGGCAATGTTAGAGTCATCGAAAGTTTTT
 AATCAGGCAAGTGTATGATCAGTTTTTTGTTTTGTTTTTTTGTGACGGAGTCT
 GGCTATATCGCCAGGCTGGAAATGCAATGGCGCGATCTCTCCTCACTGCAACCTCCGCCCTC
 CCGGGTTCAAGCGATTCTCCTGCCTCAGCACCTCGAGTAGCTGGGACTACAGGCGCGCGC
 CACCACGCCCAGATGAGAAAACCTAGGGCACAGAGAGGTGAAATAAGGAATGAAATATTAC
 TGGATGAAATAATTATACTGATCCTAGCCTTATTTTAAAGGGCCCCCAGAAATGCTCTTC
 TCCAAGACTTAGAAAAATCAAATGCTATCCATCCCTCAAAGATACATCCCTCCAGCCCCCTC
 CTTGAAGTCTCCTGCAGCTAGAAAGTAGCATCTATCCCCATCTGGACTCCTTAGC
 ACTTTGCTGTACCTGCCTTAGTCCCCCATCTCTTTACACACACACACACACACACAC

ACACACAGAGTATTATAATACTATTACCTATTATATTACATTATATTATTGTTCTTCTAC
 AGGAACTACAAGGAATTCTACTGTGAATTCCTTAAGGGTGGGATACATTTTACAAATCATC
 ACTGTTTCTCTCATTAGTTGTGGAAATCAAAAAAATGAGTGCACGATGAAAAAATTATTCAA
 ACACATACATGCTTTCATTCCAAATTGAAAAATGTAAAAATATTACTGTTAAAAACAATATTA
 AAGGACATGTAATACAAATTCAAAATATTCTACCATGCACGATTGTTTTCACTTCTTTGTG
 CCTCTTTCACCATTTATTGTGTGCATACTCAGCATATTGTGTTGTGTAAACAGGATGTTA
 AAAATAGAAATTTGTGGCTGTAGTTTTTGTCTAAAGCTCTACTAGACTGGAACAGAAACAAT
 AGGTGGTAAATACTCCATTCCCTCAGTGAGATGCTACCTACCTATGCAGAGCTGGAAAAAGA
 TTTTGCAACTGAAAACCCAAATCCTTTCTGAGATATAAAAGAACAGAGAGTCTGGAAAGT
 GATTCTTCGGAGAAATTCAATTTCTTATTCCAGAGAAAGAACTTCAAGCTCAGAAATATT
 GGCTACTACCTGTGATAAACATTTAAATTAATTGGAACGAGAGAGTTTTTTATACTAAATT
 TGAAAGAAGCAATTTTTTTTATCAAAGACCAACCCGAAGATTTCATGACCACCTACCTGTGA
 TTTTCATCCATGTTGCAAGGAAAAATCCTGCCCCCACTTCCCCCAATAGGAAGAGTGGTGGGGG
 AATAATAAAGTTGTTTATGATGCATCCTGTGGACTTCTACTCGACAGAATGGTAACATT
 CAACTGTAATCAGGATGAGCAAGCCACTTTATTAAACTAAAAATTTTAAACTTTAATAAAT
 ATAATACAAACTGAAAAAT

SEQ.ID.NO. 428 HG1014988N0 PLT00014330N10.contig
 CTATTTAGGTGACACTATAGAAACCAAGAGGAAGCAAGATGGCTGCCCTTTAGGATTTGT
 TAGAAAGGAGACCCGACTGCAACTGCTGGATTGCTGCAAGGCTGAGGACGAGAACGAGG
 TCAGAGCGCTTCTCTTATGCCCGCGAAACTCTCCCTTTCTTCTCCCCCTTCGCTTTTCTCG
 GGCTTCCAGGGACTGGGAGCAAACCCCTGTAGTGTCACCCAAAATACCAAGAGGGAAGA

GCTGGGACAGGAATCCTTTGAACTTGGGAGGCAGAGGTTGCAGTGAGCTGAAATCACA
CTACTGCCCCCAGCCTGGGCAA

SEQ.ID.NO. 429 HG1014990N0 PLT00014330N12.contig
CACCAGTGCCCTCCCTGTGAGGGCCTGGCCAGCTTGGAGCAACACCTGCCTACCTGATAGT
ATCGGTCAAGTTCTCGAGCTCCTCTCTAGAGGACCGATGGTATCAGCCAGGGCAGCTCC
TCTCTAGAGCTGTGTCTTAGTTGATTATATCTGTTTGCAGATCCTCCCTACTGCAATCTG
TTCTATGAAGTGTGCCATATTAGTAGTCTTCCTTAAATATGCCTTTAATGACTGATTCT
TATTTCTCTCTCTTTTTTTTTTTTTTATTTGAGACGAAGTTTCGCTCCAGCTTGGGC
AACAGAGTGAGACCCCTGTTCCAAAAACAACAAAACCTGGTTTCAAATATATTGCTACAT
TGGAAATTCCTCCATAAAGGAATTAATTGTTGATAACAATGCCCTCAGTATACTATTGTCA
ATACCTATAACAATGCAACAACAACAAAAAGAAAATTTAAATAGCTGGACCACATTAAAAG
TTTTATTGGGCTAAAATCAAAAACATCTGACAAAGTTTTTGACTTGTCTTTCTCATGATTT
AATGTCCCAGCAGCAACTGGGAACCTAATTCTGACCAACAAGGAGGAATTGTTTGGCT
AATTGGAAAGTGACAGAAACCTTGGGAGAAAGCAACCGCATACCAAGGATAATTATAAGAA
GGTAACATGGGGCCCTAAAGGATAGTGTCAAGAAACCTGAAACCCCAAATGAGGTGAGAC
TTGCAAAAATGCTAAGGAAGGACAAAAACAGGCCTTTTAAAATTCTGTTTCAGAACCCAG
AAGATGATCAAGAAAGGGAAAATGCCCACTAGCAGCTTCTACTGTAAAGTTAGAGCTGACA
GAGAAAGAACTCCTTAACCTACTATTGTCTTCATTCTCTACAAAGGAAACTAGAGAAGTGG
GTTGACGTAATAGAAAGAACATGTGTTTGTGGGGCCAGGCAACCTGGGTTTAATTCTGT
TTCAACACTGCTTAACAAAATTTATGGGAGGCTATTGTTTGGATTGGGCTCCTGCACCA
GGCCCCCTCCGGACCAAAACCAAAATGGAGTCACTCATATAAACTCCAGGTCACTGAACC

AAAACTAAGTTGTTTATCTGACCTTCCAAGAAATCAGGAGGGAGAAAAACCAAAATCTC
 CAAACAGGCCAGTTTTTAATCAGCGTGATAAGGAAGTCCTCTCTTTTTTAAACCTATAAAG
 AAAGTAACTTTTTTGAAATGATCAATACACTTTGTATTCTTAGTTCTGCTTCTTTAGCC
 CCTTCTGCCCTATAAAGCCCCACTTCCTCTGCTCAACTTACTGAAGCAGTATTCCAATTTA
 TAGAATGAGATGCTGCCCAATTCTGGAATCACTAATAAAGCCAATTAGATCTTTACATT
 TGTGAAATTTTGTCTTTGACAACATTAAGTTATATTCTGGGTATCCTTTTCCCC
 ATCTGTAAATGGACATAGCGATATCCTTCCCCATCAGATTTTCTCATTAATAGAAGTAA
 TACATTCAAAACACTGAGCCAGGCTCACACCAAGTGAGCAACTTGTTAATATTACTCAGAA
 GCATATTATAGATATTGACAGAAAGCAATACTGTTCTTAATTAGGGGAAGAAATTTTAAG
 AGAGTACCTAAGAGTTTGAATTTAGATTATAACATTTGCTCTGAGTATTTACATTACAG
 CCTTTGGGGGAAAGTACAAATGAGATCTGAGAACAGTGTACTCATCTTTTGAGGAATT
 ATGGAAAACGTAATAGAACACTAAACATGGGAAACATCGGCCCTTCAGGTTGAAAAGTGG
 AAATCTCAATCCCTGAATTTTTTTTTTTTTTTTACTAAGTAACTTTTTTTGCCCCATTGGT
 GTCATTTTAA

SEQ.ID.NO. 430 HG1014992N0 PLT00014330N13.contig
 CTGGGTATTAGGTGTCCAAGAAAGCTCAGATCTGCCGACCGCAGGGTCCGAACGCA
 GCCCGGGATCAGGAAGAAAGCAGGTCGTTTGCAATCTACCGGAGCCTAAGGGCAGAAAGGC
 AGGGGTCCGAATGCAGCCCCGGGATCACAAGAAAGCAGGTCGTTTGCAATCTACCAGAGC
 CTAAGGGCAGAAAGCAGGTATTATAATTTGCTTCATAATTCTGAGACTCCTCCACTTGAA
 CAGAAAATCTAAGAAAACAATCATGGCAATGAGGATAGCACTGGAAACACACATCACTATGA
 AACACCATGCTCAATACACTGTTTAAACTGTCACTGTGAATGAAAGCAATTGTATTGTGT

TGAGGAATATGGACCTAAGTCTTATACCTTTATAGTAGCATTTTGCACCTTGAGGATAAAG
 AGAATATGGAAATATTAAGAAATATTTTAGCTTTGTAGATTTTTCCTCCAGATCAAAACA
 GAACAAACGACCTCTACACATGTATTCAGGTAGCTGAAGGCATTGACAAATCAAATCAAAT
 TGTTTCATTTTGAACAAACCGAAAGGGTGAGGCAATTTCAATAAAGAAAGGTTGGGAGACT
 TCTAGAGTGGTTGTAGAGGAAATGCCAAATCCAAACGTGCGCCATTACACACAAAAGCCACTT
 GCCTCCCAATTGGGTTATCGTATCAGGAATGAGGAACGAAAGATTGGCAGAGCCTGAGGAA
 CTGGAAGAAAGAGGTTATCGGTAAACACATTTCTTGTCTGAACCAAGAACTTAAAAATTAA
 GTGCAGAGCAAGCAGAAAGTATTGACTGAAGCTCAAAAGTGAGTTATATGTTTGGATG
 GGGCATTAGATAAAAAATATGAGAGCTGGATTTTATATAAAATTTGAGTATAGAAAGGTCA
 TTTTCTTTAAACGCCACTTTTTTTAAATTTGGGCAGGTTAAATCTGTCTGCATAAAA
 TGAATAGAAAGAGAACTCTTGGTGAAGGGAAAACAATGCTAGAGAAAAAATAAAATAAA
 ATGTATGATAAGTACATGCAAAATCAATGATATATATTTACAAAATATGGTACTGTTTAT
 ATACCCAGGAATAAACCTGAAGAAAGAAAAGAAAATAGAAAATAAAGACACATAGGAGAG
 AAAATGGTTAAAACTCACTCGACACTGACGAGGTCCCAATCCATGTCTCTGGGAAGTTT
 GTTCTTACTGTTTGAATTATAGCCTGTTCTACATAGTCAGCATTCCTCCCTCTGATTCCA
 CTAGCTAAAAAAAAGGAAGAAAGTACCCTTATTCAGACCTTGGACTCTGCAAAATTATAG
 TCTATATATTATCAGGGTGAAGGTCTCCATAAAACACCTCACTTTCCTGTATCACAGTC
 ACTTACTATGACTGTTTTCCTGTATCACCATACATTGCCTGTTCCCTACACAGTCAGGT
 ACCACATAATGACACTGTGATCAATGACAGACCATTAGACTACAGTGGTCCCATAAAGT
 TAGAATGGAGCTGAAAAATTCCTACCACTTATTGACATCAAAGCCATCATAAAGTTGTAG
 TGCAATGCATTGGGTTTGTGTTAATGCTGGGTAAACAAACCTACTGCACTGCC
 AGTCCTATAAAAAATATAGCAGACACAAATTTTGTACAGTACATAATACTTGATAATGATGA

CTATGTTACCGGTTTTTGTGCTTACTATACTATACTGTTTACTGTTTAGAGTGTAAGTGT
 CTTACTTATCAAAAAAATTAACTGCAAAACAGCTTCAGGCAGGTCCCTTCAGGAGATGTTT
 CAGAAAGGCATTGCTATCATAGATGATGACAGCTCCATGCATGTTATTGCCCCAGAAAG
 AACTTTTCAGCGGTACAAAGATGTGGAGGTAGAAGACAGTGATATTTGATGATCCTGACCTTG
 TATAGGCCCTAGATTAAATGTGTGTTTGCATCTTCGTTTTTAAACAAAAAGTATAAAAGT
 TAAAGAAAAATTTAAAAATAGGAAAAAGCTTATAGAATAAGAAATGTAAACAAAAAATAT
 ATTTGTACCCTGACTCACTCAGAGTAACTTTTAACTTGCAAGCTCCATTCATGGGAAGT
 GCTCTATATAAAGTATAACTTTTTTTTATCTTGTATACCAACATGTTCACTGTGCCTTTT
 CCATGTTTAGCAATGATTAGATACACAAATAGTTATCATCGTGGTACAACTGCCTACAGT
 ATTCAGTATAGTCACATGCTGCGGAGGTTTGTAGCCTAGGAAAAAATGGGCTGTACACAT
 AGCATAGGTGTTGTCAGGCTACAGCATCTAGGTTTGGCTAAGTACACTCTATGATGTTT
 GCACAACAATATCACCTAACAAATGCATTTTTTCAGAAACAAATTTCTCCATCACTAAGTGAAG
 CATGACTGTATTTCTGACATTGTATTTTATCAAA

SEQ.ID.NO. 431 HG1014994N0 PLT00014330N22.contig
 TGCTGGGCAGTCTACGCTAATCATCTTATTCAATTCCTCAAAAAACATTCGGAGATGCAGGT
 CTCATAGACCCATTTACAAGCAAGGAGATAAGAGTTCGAGATGGTGTACATCTCATCA
 AAGGAAAGGCAAAATGCTGCAGAGCCCGGATTCTAATCTGGGTGTGGCTAGTTTGA AAACC
 TGAGTTCTTTTCCACTTTGGAGGAAAAGTGATTAAACGGACAACCTCCCCCTTCACACCTGT
 ATCCATGTCCGATGCTTTTCCAGACAGATTTTGGCTGCAGAGCTGGAAATCACAAAGAT
 CAT AATGACAGCTCTAACAGCCGTGCCCTAGGCTGATGAAAAATGCCCTTTGAACAGCCCCGG
 GCTCACAGGCATTATCTTTTGCTAATCCTATAATTGGCCTGACAAGTGGTGAAAGCATAAA

TGATTTCCGATAATGTGCTTTTGTGTTTAAAGACTGGGATCTGGGGTGAAATTCCTCTTTT
GTCTTTCAAGTCCATTGTAGTTACTTTTTACAAAAAATTTGTTTGTGTTTGTGTCATGT
AGACATTTTAAAGCAATTTGTAGTCCAAATTTGTTAAACTTTCATTTATGATTTGTTTAA
TAAATTATAGAGAAACTACAGAACTAACTTACTCAATTTTCTTGTAGTCTATTTATTTA
TTTTACATTTTGGCTGGGCATGGTAGCTCATGCCTGTAAATCCCAGCACATTTGGGAGGCCGA
GGTGGAGGATCACTTGAACCCAGGAGTTCGAGAATAGCCTAGGCAACATAGGGAGATCC
CGTCTCAACAAATAATTTTAAAAATTATCCAGCGTAGTGGCATGCGTCTGTGTGCCCAGC
TACTCTGGAGGCTCAGATAGGAGGATCACCTGAGCCCAGGAGAAATAAGGCTGCAGTGAGC
CATGATTGTGCTGCTGCACCTCCAGCTTAAGTGCACAGAGTGAGACCCCTGTCTCAATAAATT
TATTTATTTATTTATTTATTTTAAAGTGTGTGGCCCATTAGTGATACATACAAA
TTCAGTTTTCCTTTTCTTTTTTTTTTTTTTTTTTTCAGGTAGCCTAGCTAATCATA
TCTCTATTGATTGACCAACTGAAGGACGAGTTTACTACTGATGGGTCCCCAGTCATTCCC
ACTGGTATTCAGACTAAAAGCAACATTGGTGGCAGGAAAGTCCACAGGAAAGGCTCTATT
TCAGGAAGAAAGCCAGAAACCCCTGGAGAGAAATCCTAGTTCTGCCCCCTAACCTGCAGGATGG
CCTTAGATGGATCTCTTTTGTACCTTACTTTTGAGTTTATAGCTCTACCACTAGGGAGAG
AGGAAGAGGAACAAAGTAATCTCTGGCACTGATAATTATCTTTTGCATAACCCTTGAATTC
TTGTGATAAAAATAAATGTATAGACCGTGAGCTACTGTGCCCCAGCCTTACGGACATCCTT
TTGAATTAATCTTTTTCACCTCATAGAATATGAATACATTTATTTAGACTTTTCTAGAACT
TTCCTGTTTTCATGTCTTTGCTTCATCTGGAAATTGGCTTAACACCCCTTTTATAAAGTTTG
TGTTTGTAAAAATTTCCATTGTGACATCAATACGCAATATATTTTGTAAATATAGGAGTTTC
TAT

SEQ.ID.NO. 432 HG1014996N0 PLT00014330O03.contig

ATCTCAATGAACCTTCATGGCTCAGTTGTAAATAATAATTATATTCTCTAATAATGACTAA
AGAAAGTTAATAATTTATAGAGCACCTTAATGCCTGGCATATAACAAAGTATTATATGTGTTCA
TTAAAGTAATGCCATAAATAAAATAAAAGGATAAATAATTGGTTTAACCAGAAATTTCAA
TATGACTATAATTGGGAGAAATAAGAGGAGAAATTGTGGATTGTGGTGTGTCCTTTGTGTTT
CTAAGCGTGGGGTGGTTATAAGATCATTAAGTGCTTATGTTTGAATGAAGGTAATG
AACAGATACTAAGTAAAATTAAAGCAAGAAGTATAAGTATAAACAGCATAACATACCCCTTC
TATGAATCCAGTCAGTGGTCTCTAAAGGGAATTATTAGTAAACCAGCAATTAGGTGAAACA
TATAGAAGTCTCTTTAATGCATTCTCTGTGGAAATCTCTCTAGTCATTTCTGTAGAAAGTC
TCTCTCACCTGTTCCGGCAGAAAGTCTATTTGAAAAAGAAAACTAGCCAAAAATTTTAT
TTGACCTTGCTTAATTCCTTTGGTTACTGTAATTTCTATAACCTCATCATTTATTGATAGTCT
TTACTTTTACCTCTTTTCTCTGTCTCTCTCAACTTGAGGCAATTGATTTTCTCTCTT
TCTTTTCCAAATCATCTACACTCTTGAAGCCTGGAAGATGCCCTTGTGTGGCAGAAAGGGG
TGATCCCTTTCCCTTCCCATTTGTAAAGGTCATGACCAACTTGTAAACAAGACACAAAGTTAAA
AAAGAAAAGCATAAAAAATTATCTGTTTATAGTTTTATGTGACATGGGAGCTTTTAGAAG
GAAGACCCATATATACAGGAAAAAGATATATATTTTATGCTTAGGGTCAATGAATAATGA
ACAGCTGTGTAGACTGTGCAGAACTGATTGGGCAAAAAACATATGATATAATGCTAACAA
ACTGAGTGAGGAAACTCAGCAAGGCCGTGTCCAGATTCTTCTGGCCTCTTTTGCAGCATCA
ATTCTCCACGTATTGGGCAGAGCCCCCTTCTGGAACGGGTCTTAGGATGGACAAATCAGA
CAAGGGAGGTCAGAGAATTCTTTATGACCAGTTTTTAAACAGAAATGTGGGGCGAAGTTA
GAGTAATAATTTTAGGTTTTATGGCTGGCTTAGGGGAAAAGGGATTTTGTTTTTATGA
CCCACTCGGGGAAGAGGGATTCAATTTCTATGGCTTGCTTCAGGAGAGAAATGAGGGATG

AGAGACAGAGGGTAGGAGAAAGGTCAGAGAGAGAAAACCTTACTTCTGAGGTCTTCACITTT
GGGGTATAATTTTCTGAGCCCCACACTCTCTTAAATGTCACTAAATGGACTTTTACTCACT
CTTACCTATTTTATTTTGAATTTAGAAAACAGGGTTTCACCTATGTTGTATAGGCTTGTC
TCGAGCTACTAGCCTCAGGTGATCCTCCTGCCCTCAACCTTCCAAGTGGTATTACAGGTGT
GAGCCGCTGCACTCAGCACACACATTTTATTTTATCTTTGAAAAGTTTTCAAAACCTATA
GAAAGCGAAAAGAATAATACAATGAATACCCACCTGTCACTTAATTTACCAAATTGCT
AACATTTTGCCATGTTTGCTGCCCACATCCCATGTGTGCACACGTGAGTGTTCACACAGA
GCAATTTGAAGGAAAGTTGCAITTTATCATGACTTTTAACTAAGTGCTTCACTCTAAA

SEQ.ID.NO. 433 HG1014997N0 PLI00014330007.contig

ATCCTGGCCGCGCAGCTACCTTACGTACGCCGCTCTCAAATCACAGGCTCTGGCCAG
TTCTTTTCAGCAGCTTTATGAGATTTTCTCAAATCTCAACAATCCGGTATATTTTCAAAA
GTACATACTGAGCATCAAGGTGGTGCAGCGAACTCCAGATCTGCAGCTCCAGCAGTCT
TGCTTGAGCGTCCACTCAGGATGGAGGTGAACTCGCAAAGTAACTGAGATTGGTTGGGCA
TGACATAATGGTGGATCTAGGGTTGAATGACCTCCCCAGGAAACGCAAGGCTTGAAGTGGAG
GCTGAGAAATGCAATCCGAGTACAGCCATAGGCCGGTCAAAGGGTCTTGCTGTGTCACCCCA
GGCTTAGTGCAGTGGCGTGGCGTGTTCATGGCTTACTGCAACCTCAAACCTCTGGACTCC
GAACAGCTGGGACCAAAGCATGCACAGCCTCACAGGCTAATTTTAAATTTTTTTTTTT
TTTTTTTGTAAAGATGGGTCTTGCTAAGTTACCCAGGCTAGTTTGGCCTCAAGCGACCC
TCTCACTTTGGCCTCCCAAAGTGCTAGGATTATAGCGGGAGCGCCAGTGCCCTGGCTGGGA
CGTTTATAAAAAGGAGCAATGCACAAAGAGGCAACAAAGGCAAGTAAAGAGGAGGCTGGTG
GGCTTGCAGAAACGCCCTGCAGTACCTATGCTGCGGTTCTCCCACTTCACTGATAACAAA

TCCTCAGTATCCACTGATACAAATCCTGGGACTGCTGGGCCAGTCCCAGGAATGACACTG
GGACAGACTCCCTGGAGTGGATCAGAAAGCTGCAGTTTACCAGGGTCCTGCTGTTGA
GGCAGCCGGCTCAAGGACCAAGCCCAAATAAGACACTGAGAGGAAGGCTGCAGGCCAGGG
TGCAGGAGTAATGACCAGGATCAGTGCTCCAACCTGGCCCCACTGTGTCAATTAAAAGGCTC
TGCAATGGAATCCATCCTGGTCCAGGCAGGGGAGTGTGCCCTCCAGCAAAGGACCAGGCGT
TTCACTGCATACACAGCAGGTGTTGCTTAGTGAAAGACAAAACCCAGAGCCAGAGGTACT
GTAAACATTCTTCAATTATGCATAAAGTGACAGTCTGTTAGCTGATGTGATCCTGCAAAATGA
AGGCTGTCTACCCACCAAGGGTTTCTCCCTCTGCTAGTAAAACTGAGACTTTGTCACTGG
AGGCTCAGGTAAGTCCCCACAGTGCTCATGCCCTTCCAAAACACGCCCATTTCCAGTGAT
GATGTTTGTGCCACAGGCCAAGTACTGAACATCATGACACTCATAGACACTCTTCACTG
AAGACGTGGATAGACCCCTGGGGCCACCAATCTCTCTCTCTGAGTGTGGAGGGGCCCT
GACCAGGGCAGGCCCATGTCTGGAAAGCAGGGTTAGCCCTGGCTCATGGGGGACCCCTG
CTGGGCTTAGCCATTGCCACCTAGGGATCAGGCAGCCTTGTTTCTTCTTCTTCTCCAG
GAGAAAGCCCCAGAGGGACACTTTTCTAATGGTGAGATGGAAACACGGGAGCACCTGCCAACA
GGCCTTCAGTTATGGTCAGAGGCCCAACCAAGTCTGTGTGTAGTACCTAATCAAAACATT
TAGATTTAGGCTTTAGTTATTGCCCTGTTTTTACTTGAAATTCTTGGCTTAGAATTATTCC
CAATGGGGCCAGGCACAGTGGCTCATGGGAACAAATTAAAAATAATCAGATTAGCATGTC
AAAAGATCTCACGTTATAGAAAGTGAAGATCTAAGACCAAATAAATACTGATAAAAAATA
CAGGCCAGCAGAAAGCTTGGATGATTTAAATAAATTAAATTGGGAGTTGGGGAGTAGGGAA
GAAGGAGAGGAACATCTCTTTAGCAATTTGTTCATTCTAACATTCACATGTTTGTCTCTCC
ATGGAGTTGTTTCACTATGAACAGGTTCACTAAGGAGCCAGTTATGTCCATATGAAATGA
GAATA

SEQ.ID.NO. 434 HG1015005N0 PLT00014330O18.contig

GGCCCTTGAGAGCTCCAGGCCATCAGAGATGGTGCAGGAATGGGGTAGGAGACAGAGGA
AGCCCTGCAGGAAGCCAGCAAAAACCTCACCTACCTTGAGAACTCAGCAATGCCGGAAGA
GATTGAGATTCTAGAGAAATGGAGGAGCCGGGAACAAAGGAGAAAGGATGGGATCCAGGAA
CAGGCCAAGAGAACCAAGCTCGATGCTTCTCCTCTCCCTTCTAGAAAATTCTGTTCAGAA
GCCATTAGGCTTCAGGCTCGAGACAGAGAGGAACCGCCTCTGTGTTTTTCACTGTCGTGT
CTCAGTGTGGGAGTGTA CTCCATGTGGTCAAGTGTGATCAATTTAACTGTGCCTGAGG
GTTCTGAGAAGACCTCACAAATGAAGATGGAAGAAATCCCAAGAGACCAGTGAAGCAAGA
AATTATCAAAGCTGCGGTATTTACGAAGTCACTTTGTGCCAGTGCCCCAGCAAGCTGCCA
CCATGGCTGCCAAGCCCTGTCAACCCTTGCTGATCAGTACAGAGAGGGGCTCGCTGCGCTG
GGTAGAGGAAGCAGGAAGTTTGCTGAACCTCAGACGGAAGAAAAGGATTCTCTGGAGAA
TCCCAGCAGGTTTCTGATTGGACATGGGTCCATGTCCTTGTGAAGTCTCCTCCTCTGAA
CCTGCCCTGCTCTGTGACTTGCTTTTGACCACCAGCATGCTGCAGAAAGTGATGCCGTGTG
ACTTCTGAGTGGGTCA TAGGAAGCCTTGCAAGTGTGGCCTGGGTCTTCAGGAATGCTTGT
TGAGGAAAACCACTGTCA TGTAAAGATATGGCTACCCCTGAGACTGTTCATGGTGTGAGGA
CTCCTAAATGAGTCACATGTGGATGGCCACCTATTCAAGGCTCAGTTCTTCCAGCCATTG
CGACTGTGGCCCACGCATGTGGGGGAAGAGGCCATTTTTACGCCCCATCCCCAGTTGAGTC
TTTAGATGATCCCTTCCCTGGCCCTCCAGCTGACTGTAAACCAGACTCCAGGCAGGAACCCAC
ATAGGAACTGAACCCAGTCAACCCACAGGCTGGGGAGGATGATAGTCCATTGTTGTATAA
AGCCACTGCATTTGGAATGGTCTGTTGTATACTACCTAATCAGAAACAAGACTCAATCCTT
TCTTCATAGCTGGTAGTCTTAAGCATTTTATTAAATTCATATTTTAGTCACATTATGAATTT

TCTATCACAATAAGCAATTTATTGAGCACTTTTATAGAGGGACTATGCTAGGCTTGAAGA
 ATAAAGACACATAGGACACGGCCCTGCTTATGTTCAAGTAGCTGTGTCTAAACGAGGA
 CGCACACATAAGGTGAAGACCCCATGAGATGGAATACACGGGGATAACAAGCCTCACAGT
 GGACTGGTCGCTTTACCAGGGGTGAGGCGATTCTGCCCTATAAGGTAGGGAAGGACTTTCT
 GAAAGGGTGACACTGAGTTCACCTTGAAAAATGCCTGAGAAATTCACCCATAGACACAGT
 GGAAGAGGGGTGTTCCGAGGAAGGAGATCAATGTGCAGAGAGACAGAAATCCCGGTTTCG
 CTTATTCCCAACAAGGCAGAGAAACAACCTCTCCCAGTGATGCCCTCCCTTGTTCCCGTC
 CCCAATCTTTTGATAATAGTGGGAAGGTTTCAAGTGAGGGTTAAAAATGGGCTGGTTT
 TGCCACCCCTTAGTAAGTCTTGGGCAGGATTCTCAACTCTTGACACTCATTAATAATCACC
 TGGGAAGCTTTTGCAAAACTCCAGCATGGTACAACCTACTTTAGAAACTGGTGTATCCAT
 TAAAACTAAACATTTGCTTATCTTAGCAATTCCACTCCTAGGTTTATAGAGTAAACAGAA
 ACACATACGTGCCTGTGTCAAAAAACATGTATAGGCATGTTCAAGACCACACTGTTTGGC
 CAATGCTGGGATTGCAGGCATGAGCCACTGCGCCAGGCCCATGAATGAATCTTAAAAATA
 AAATAGTGAAATCAAGGAAGTCAGACACAAAGAGTACATACTCTACGGCTGCAITTTCTATA
 AATATCAAAAAGAGGCAAAACAATGTGTGTGTGTTAGAAAAATCAGGAGAGTGTACTATCC
 TGGGGAGGAAGGAGAACCTGAAGGAACTTGAGGGGGGCTTTGGCCTGCTGGTAGAGCT
 TTTTTTTTTTTGTCGATCTAGGTGGTAGTTACATACGGATATTCACGTTGTAAAAAGT
 TTGCCAAGGCTTTAATGTAATTATCTGATGCATATTATAATTAAAAATGG

SEQ.ID.NO. 435 HG1014999N0 PLT00014330P07.contig
 GGGGAGGGGCTATGGTTGTTGCTCAGAAAAAAGCCATCTCATCCCCCCTCCCTGGCATC
 TTTCTGATCTGCTGTTCTGTAGGTGGAATGTTCCCACTTTGTTTGGCCTGTTTCAGACG

TGTTTGTCTTCCATTATTTTTCCGGCTGGATTGCCAGCAGAAATGCAGAGCAGCCTGGAG
AAATGAACACAGAGCTGACAGGTCTCTTAATGAAAGCTTGTTTGCACACTCTGCCTCTCCT
GTGATCGCTCTGTGCCTGAGAGCCTCTAGGTGTGGATTTTTACATGAGCGTTTGTCTTTC
TTTTGATTAAGGCCCTGTAGAAGTGGATTGGTTTTTTGTAAACCAAGCATTTCTCATCCCTC
ACTCTTGTCTTCTCCTCTCCACACCTTCTCTGCCCCACCTTCTTCTCCCCCTCCCCGTTCC
TTCTCTTGTCTCCACAGCTTTTGGCAATCTTAAGCAAGCTAAATTAGTTTGGCCCATCGTT
GCTCCTCTTCTGTGACCTCCCTCCCAACATGCACCACCATCACGATGCAAAAGTGCTAGA
AAAGGGCTCTGCTATCAATTATTCACATGTGGAGTAAATTTTTTGATGTGTTATTCCTA
CCGGGAGCCTTTGCACCCCGTTTCACGATATGGCATGAAAAGTTCTGTCTCTCTCTTTAAG
GAATGACAAAGACCAGCTGCTTACTCAGTGGTGATCTTCTGTCTCTTTAAAATTAC
AGCATGCCCTTTGCCCTGTGGTTTTTCACTCCAAAAAATGGGGTGCCCTCTTGGCTTTG
TTGTTTTCCCAACAGCCTTGTGGATGACTTGGTACCTTCTTTTTTTTTTTTTTTT
TTTTTGAGAGTCTCACTGTCAACCCAGGGTGAGTGCAGTGCGCATGATCTCGGCTCACT
GCAACCTCCACTTCCCCGAGTTCAAGCGATTCTCCTGCCCTCAGCCTCCCGAGTAACTGGGA
CTACAGGCATGCACCAACCATGCCTGGCTAATTTTGTATTTTGTAGTAGAGACGGGTTTC
ACTATGTTGACCAAGGCTGTCTCAAACTCCCAACCTCGTGATCTGCCCTTCTCGGCCTCC
CAAAGTGTGGGATTGCAGGCGTGAGCCACTGCGCCCGCCTGTACCCCTCTCTTTTTG
AAAAAGAGAGTACCTTCTTCCCTGTCTTTCATTCTGAAGCCTCCATTAAACACTCCTC
CTCCTTAGCAATAGTGAGGGTCCCTTCAGACTAGATAGGATGTCAAAAGGCCAAAGCATT
GCCCCCAACACACTCCTTTAATGTAGTCCCTTCTTTCCAGGTGTCTTCCCTGAAGAGAGAA
AAATTTGTAGATCAGAGAGATGCAATTGGGAGCTCCGTAGTGAATCTAGTTAGAAATGAG
TAGCAGAAAAAATAGGAAAGACCCGTCTCTCAATATTATGGTTGCTTCCCAAAACAAAC

CAGTTGATACAGCATGAAGGCAATTTAAATGATAGATATGGACATGCTGCCTTACATTT
 CATATCACCTCTTTACTAAATGAGAAATTTTTTGGATGCATAGGAAAGTTTGCATAATAT
 TGAAAAGTTTGACTTTCTGGAATCTATGTGAGCTCAGGAATATATTGGGGATCTTAGCA
 GAGGCCAGTGTAGCCTATTAAATCTAAACAATTGGCTTTGTTTTTAAACCAGCTGCT
 ATTTTGTAGTTTCATTTAAAGAAAAAGTACCTTCTATTTCATTTGCAGAAAGTTGCTTTATT
 GGATTCTATGGTCAATTGTATATCATAACTATA

SEQ.ID.NO. 436 HG101500IN0 PLT00014330P09.contig

AGAGCTGGGGCAGGATGCACACCCCACTGCCCCGAGTCTCAGTCTTTCAGCTGCCTTGGC
 CTCCAATCTGAGCTACTCCCCTGACAAGGGGATCCACCATAGGGCACCTGGGCTGAGT
 TCCAGGGAAATTGGCGGGGATAGCCTGGGCTCTTTGGACCTAGTTTCAGGCCCTGCTTTG
 GGACTGTCCTAGCATCCCAATTGGATGTGTAGGATATACCTTGTCTGATTGTCTGGGAGAATC
 CTTCAATTGAATTAAAGAAATTAGGTCTAGAAATCCTCAAAACCCAAACCAAGAAATCTGGA
 AATCCTCAAATTCAACCACAAAATCCTAGATCTGGCACAGCCTTCAAGGATTCTTCCAGT
 CCGGTCTCTGCCCCAGGGGATGAACACCTCAGCCAATCCATGACAGCTGTTTTTCCCTCCT
 GGCCTTTAAAATCCTCCCTCCTCAGCCTTCACTCCTACTCACCCACCTCAACTTAATTTCT
 CATTCCCCACCCAGGAAGAACCTGTATGCCCCCTGCATCCCACCTGCACCTCCGGATGCCT
 TTTCCCCCAAGACTGGAATGGCCTCAGGCCCCCCCATCTCCAGGTGAGTGGCCCTCACCTAC
 AGACTCAACATATGGCCTTTGGCTCTTCCCACCTTCCAAGAGTCTTGGAAAGGGATGGGTCTG
 AGCAAGCAGAGGAAAGGAAGATGTGAGTTCCCAAAATGCTCCTCACCTTTTTTCTTCTGAG
 TGGGCTCCTTCTCACTGGCAATTGGAGGGCTTGGCGGCAGCATGGTCTCCACCCCTGGGA
 GACTCCGTCCCCTGCTCTCCTAGGTGTCAAGATGCAGAGGCCCTCTTGTAGCCCTCACCCAG

AACTGCCCGGGGCACGGCATGAACCGAGCCCTTCAGCTGCCACTCTGACTGCTGCAGTGG
CAGCTTGGGGGTGCGGCCATGGACACCCGGCAGCAGCAGCATACAGGAAGCCCTGCCAC
GTGACCTACTCTTACAGCAATCGCAGCCCCCTGCCGGCCCCCTAGGGAGGAAGGAAGTCCAA
GCTTCAGTCTCCAGAGATTCTGATGCAGACGGCTCTTGGGTTGAACAGTCAGAGAAAGTC
CCTAACCTTGTGTCTCATGTGATCATCATCAGTCCTTCTGTCTCCATCTTCTCCTAGAGGTG
GACAAATCAGCTCGACATTATATGGGTACCCAAAGCCCTTTATCTCACTGACCCAGAATCCA
GAAACCAAGATTGGACCATCTGGCTGCCTGTTGGATCAATGCTCCCGGTCAACTCACGT
CTTCAAGGTCGTCACTGCCATCACTACCACTACCATACCCCTGTGACCATCACCACTGC
TACCATCACTCCTATCAATAGCCATCACCATACAGCCATCACCCCTACAGCCATCAGGCC
TGCAGCCATCCCCGCTGCCATCAGCACACCTATAGCCATTGTCACTATGGTAATCACCTC
TACAGTTGACATCCTTCAGCCATCACCTCTACAGTCATCAACATTGCTGCTATCACCCCTT
CCAGCCATCATCCTTACAGCTGTCAACCTTCTTGGACATTACAGCCAAACCGCACTCATGAC
CCTCTACCCCTATAACCATCGTCCCTACAGTCATCTCTCCTACAAACCATCATCATCC
GTGCAGCCATCAGCCCTGCAGCCATCACTCCTACAGCCATCAGTACACCTAAAGCCATCA
TCACTAGAGCAGTCAACCCCTCCAGTCAGTACCCCTCAGCCATCATGAGTATATTTATCAC
CACCACAGCTGAAAGCACTTACAGCCAACTCCACAACCTCTAGCAAAGTATCCAAACAGGCA
TTTATCAAAACACCTATGCAGTATATTCAATTTGTGGGTGTCTGTGGCAGATTTTAATAA
CTTTAATAACTTCTTGGTATTACAAAGCAACTTGTATTGTCTATACAATTTACAAACA
CATAAACAAAAAAATTTCAAAATTTCCCATCATTTGTCTCTACACTGTTTTAAAGCATCT
ATCTACTCTGTCTATAAACATTTTTTTTCTTTTTTACAAAAATCATACTGCTCAGAAAA
CAGATAATTGTTTGGGAGAAAGGGAGGGGAGAGAAAGGAAGGGAGACATCACACAGGG
TAACAGGAAGCTTTTAGGAGTGATGAATA TGTTCACTATCTTGA CTGTGTGATGATTC

ATGGGGGAATACACATGTCAAAACGTATCAAAATTGTATACTTCAAATATGTGCAGTTTAT
TGTAGCTCCATTATACCTCAATAAAGCTGTTTAAAAATTTTTT

SEQ.ID.NO. 437 HG1015003N0 PLT00014330P15.contig
CCTTGGGTCTACACGGTTTCCCTCCCAAGCCAGGCAGCGCATCTCTCAAGAATAGGGAA
AGGTCTCCATCAACTTTGCCGTACGGAGCAGTGACCGGGGGCGGGGCACACAC
CTGCCTGCACGCGGGCCCGCTCCCCACCCGGTCCCCGTCCCTCAGCGCTCCCCGGGCG
CGAGATTCTAAGGACACGGCGCTCATTTCCGGGCTCTAAATCTCCACTCGCAAAGCTGT
GCAGCGGGGAAAGGGCAGTGGGGACACTTCCAACGGGGAAGCCAGCGGGAGCGAGACAC
TTATTTCAATGGCCTTGTTACAGTACATTTTATTCCCAAGCACTCGCGGCCGTTCCAG
GTCTCTGCCCGTTAGGAAGTGTAATCGACGCCAGACCCCTTCCCTGGTGCACACAAGAG
ACCAATCGTGCGCATGTGTCTAATCCTCCCGCAGCCTAGAGACAGGCAGGCGGGCGAGGC
AGCCGGGCGCCGACGTCCTCTGCTGAAGGAGTGGAGAAGACCGACCGACCCGAGTGGTTG
ATTTATGCATGCGTTTTCGTTAACTAGGACTGAAAAGTGGCTCCTGGCTCTTCTGGGAAAA
TGGAGAACGTGGTTGTTATCAATGGATACCAATGACTTCCCTGAAATGCTCAACTTGTACA
AACAAAAAATGAAAAA ACTTTCTCCCTTAATGTAGAAACACATGGCATTTTAAAAAGTAGG
TTTTTCAATTATGTAAAGATGTTTAAATTTTAGGATTAATGTTTGTGAAAAGATTCGAG
GAAAAAGAGCGATGTTGAGTTTGGCCGGAGGCAACATATGGATTTTACTATAGGCTAG
AACATTTAGTACTTACTTCTAAGGATGTTTCTAATTAGTAAAAATGTATAAAGATTGACAA
TTTTTAAAGTGACTCCTGCAAGAGTTATATTTAGTAAAGTTCAAAAGATGATGTTTGT
TATTACGGTGAGCAGTATTTTCCAGGTTCTCATTTAAAAATTTCTACCCCATTTGAAAC
ACAATAATAGAAAAAGGAAACTTACTTGAATGAGTTGTGGGGGAATGTAAAAACTT

GAAGGACAAAGGTAGCTCAGTCCAAACCAGATGAGCATCCAAGGGAGATTTTGTCTCATCC
TTTGCAGAAATCTCCA TCTAGAGCCCGTGTAAACCAATGGAAATTCTTACTTGTCTTTAAC
AGCACACTATTCCCTATTTCATTTCATTAAATGTGGGCTAGCCCTTTT TTTTAAATAAAA
AAACCTTCCACCCGTGCAAAAGTACCAGATGCTGGGCGAGATCAGATACTAAGGATAACATG
GTAAAAATCCCTTCCGCTGTCACTCTGGAAATGGGCCCAAACTCTGACCTTCCCAGAAAAATT
TGCATAAACTCCCTGGAATGGGCTCAAGCTCTGACCTTCTCAGAAAAATTGCAATAAACTC
TCTGATAA ACTCCTTGTAATTCTTCTCTAAATACCTGGCACAGCTTGTCTATTCCACCC
ACTTCTTACTAATTCATAACGCTGCATGGCCATTTTTTTTAGAAATGTAAAGACTGCTTTA
GGTTTGTAAATGCTTTAAATAAGTGATTTTCATTTCAGCCCTCTTTTGCAACCACATAATGTTA
AATGTGGTTGCCAGTTACTAGAGGAAATTGAGGCTCACAGACATTTCAGTTGCAGAGCCAC
ACGTTTCAAGATTGATTTGAGGCTACAGTGATTTCTATGACCCTGACTCCTTCTCCTA
TTTATATAATCATCTGTTCACAAGTCAATTATCAGTTCTTTGAAATCAGGGATTAGTTT
TGCTCAGCAACAATCTTAGTGCTTACATAAGTAACAGGTGTGATCAATAATCGTAGCTT
TTAATGATTATGTATGTGATCACTGTGCTACATTTCATGTTTATCAACGTCTTTGTAAACA
CCCTAGCATGACTTCAATAATTCTCATTTATCTTTTATTTTAAATTGTGATACAGCAAT
TTTTATGTATGTGTGCCCTCTTTTATGATCTTTATTTGCTTCTTGAAAGCAGTAGGTT
ATCTGTCATGTAAATTTCCAGTTCCTGGATTGGTTGCAT

SEQ.ID.NO. 438 HG1015007N0 PLT00014333A03.contig
TGAAGATGAGTACAAGAGAGAAAGACAGGGCTCTGAAAATACTGCCATAGGCTCAAGTTC
CAAAGTGTGAGTTACAAAAGTATAAAGGACAAACGCAAAAGGACTTTTAGCCAAAGAAAG
AACCAGAAAGGTAAACACAGAAAGCAGGAAAGAGTGTGGGCTCTGGGGCTGGCTGTATCTG

TGTTTGAATCTTGGCTTTGTCCCTTACTTCCTGGTGAACTTGGGTGGCTTTTGAGCCTGC
 ATACACCTTGACTTGAGTGTGGCTTTTCATTTGTAAAATGTTGATGATCCGTGGCCTACC
 TCAGAGGTTTGACGTGAATGAAATATATTTGGTGGAGTGGCTGGACTGTAGCGAGAAATTC
 ATCACATAATTTGTCTCTTATCAGCGATTATTAAGTAAATAAATTAAGACCATAAGTGTAAAGAG
 CATTTAAGGAAAGGCATACATTTAAAAAGTAAGCTACAAAAATAGTATAACAGTGTGA
 TTACACACACACACACCCCACTTTAAAAAGAAATACACCAAATGTTAATTTTAAAAACA
 CTTGGCTGGTAGGTAAAAATAAACTGGAACACAATCAGAGGAGAACAAATCATATCTCCTG
 AGGAGCAGCTGAAAAATGGGTCTATTTTCCATGGAGAGAAAGTGGACACGGAGGAATTTTAA
 GGGGTATTCTGGTTTGTAGGATTGCAGCGGGCTACATCTAGAATCAACGGGTGGAAGTT
 ATATGGCATAAATCTCAGATCAGTGAGTCAAGTATCAATTTGATGCTTATCTGAGCTTICA
 GAAGGAGGAATACTTTATCAAATGAGTTCCTTTATGCCCTGGGCCCTGCCACTGAATGAGTT
 ACATGCATTCTCTTATTTAAACGTTCTCTCCAACAGGAAGAGTCCTAGGGAATTTGGAGAC
 GTCGCGAAGGAATGTTGTAAGGAGAAATTCAAAGCCTAAACGGTGATAAATCACAGTATCT
 GCAATATCTTTTGAACACTCATGGGCTTTGGCCCTCCCATGTTCTAGCCTTGGCAGGACAGA
 AGGTGGTAAGAGCAGTATCTGGTAGGGGTCAAGAGTCTGGCTCATGGTGGAGGCCAAG
 CAGAAAGTGACCTCATGGGGACCGCCCTGGGTTTGTGTGCCCACTGGAGAGCTCGTTC
 CTATTGCAGGGCCATAATGTAAAGCAGAGGGTGCTTCTTTTCAGAAAAGCTCACCTATAGC
 TCCTTCAGGCCTGTACTCTTAGGAGGGGAAAAAGGAAGGGAGTACAGGTTCTGGAAATTTAA
 AGAAACAAAAAGGATTATTCTGAACATAAAGACTTTAGAGGACATGTGGGTTTCATATTT
 TATGTCCTTTTATTTCCTTTATTCCCTTTTGCTTTCCAGGCTTCAACCATATATCTGGCCCT
 CACGTCCAAAAAGCTCGAGTTGATATGTTTGTGTGCTGTTTGTTCCTTATTGTCCACAGAA
 CTTGAGGCATGGGCCACAGTACTGGGGATAAGGAAGGTGTAGGGAGTAACAGCCAGAGTA

GTTCCCTTCTCCTTTTCCCAATGGCAAGGCTGAGAAACCATTTATCAAGCATTTTATT
GACTGGGACAGCTTGAAAGCACTGTCAATTACCTGGGCCCAATTTCTCTCCATTTTCCTGG
TAATTTCTGTAGTCACTGAGGAAAGCCACATTTCTGGTTGGTCCACGGATTTCCCTTATTGC
AATCATTTAATTTGCCCTGCACTATAGGCTTAATTTCTCCTTCTCTTTTGTACCAAGAGAGG
TGCTAAAAAATCCTATCAATATCAAAATTGACTCCTAAGCTGAAGAAGAAAGTTATTTTCT
TATGTGACGTGTAAAGTGATCCTCAATATAATCCCTGAAGTTGAACCTTCAGAAATGTTTC
GTGCTGCAGCAAAATTGGGGGGAACATTTGGCAAAATCTTCCAAGAACACATAGGTCTTG
GGATGGCCACATGCAATTTGGCTATCTACATTCCTGCACATCTGTGGAGAGAAAAAAGTT
TCCTTTTAAATACGGAGTCATACGTTTATTCTTAGGCTACTCCAAAAGATGCAACAAGG
TGTTTTGTGTTTTATTTTCAAAATGGGTTTGACAGTCTTTCTGCCTACTGATTGTCACA
GCTAATTGCTCCTTCATTCGCATGAGAGAGAAAAATAGAGAGGGAACGTGATGGCCC
TTTACAGCAGCTTCTAGGAATGAATGTGTATAAAATAATCTACTGAATGGAGCACAGAAT
GAAAGCAACACATAAGAATTCAAAGGTACATTTTAATGATGGGTGCATATATATTACCA
ACTTGGGGGAGAAACATGGCACCTGCTACTCTGCCCTTCACTCAAAGTGTTAAATGAAAA
CCTCTCAAAAAATTAAACCAAGCTGTCTTATGACTGACATTTAATTTTGAAAGATGTGAT
TTTCCTCCGTATTTCCCTGTGCTCTTTACAAACCTGTTTATCAAAATGTATAGAAAGACTTTT
GTTTATGACAGTTTATGATTAAATCATTCACCAAGGACACGAGTCCTTTGTACCTATTATT
GTCAAGCATTTGTTTGGACTGTGTTTGACCCTTTGTACCGAAGTGGTCTGTGTTGGAAACC
TTTTGAAAAACAATCTACCAGGCTGAATGTGCTCATGCCACTTACTGAAAATATCATGGCTC
TGTTGCTGTGACATCCTGTTACATGGGAGAGACCAAAATTACAAATTTTGTGACAATATC
AAATCTATTGTTACGCCAGGATAAATAGATTTCATACCCAGTTAGTTGTAAAGTATTCTAGA
GTGAAATTTTAAAAAACATTTTAAGAAATGTATGACTGTCTGGTGACCAGAGTTACA

SEQ.ID.NO. 439 HG1015009N0 PLT0001433A06.contig

GGTGAGCACGAGACGTGGTGTAGGGAAAGAGGACGAGTGAGCAGCGCCTGGCTGTAGGGT
CAGAGGGCGCCTGTGTCATCCTGAGAGCCTTGCCCTTCGCAGTTTGAGATTAGGGGAAGTTC
GGGGCCTGAGAAAGGTGGCGCGATGCTGGAGCTGGGAAATTCCTCCTCGAGCCCCCTGGAC
CGAGTTGCGCTGCGCCTGCAGCAACAGGCCGGGAAAAAGTGTCGGAACGCCCTCGTGGACT
GCAGCGGGGGAAATGTCCCTTAAAAAGTGCACGAAAGTGGGGAAAGGTGTAAATTACTAT
TATCAGCATCTAGAAAGCATCATGAATTTGCTGGAGTACTTCCTAGCACTGACCTCCTTC
ATTCTGCGTTGTCTTACTGGATCTTCCATCAGCCACAATAATGGAAGTACCAATACAA
GGTCAAATCATTCCTGGATTTCATCTGGAGTTGCTTAAAAAGTTAAATCATTTGGAATTTTGG
ATGATACCTTTTCTATAATGGATTACAAATTTGATCGCTGGGAAATTCCTCACCTTAAAGAAAG
ACTCTTTTGTCTATCTGGAACCCCATGTCCACCTCTGACTTCCACTCAAAATTGCTTTTCCT
CACAGTCTGACTGCGAGAGTTGTGAAAAATTGGGATGTGCTTCTGAGGTGGGCTGTGGAA
TGCCATTACCCCTCAGGTGACTACAGATGTGTTAACACCCAGCATGTTCCGGTAGGAGACT
TTCTGGATGGGGAAGATTTCACAGGTGTGAATCATATGGCAATACAGTGAAGACAGTGATTT
ACTGCTTTTGAGGGCGTGCATGTATATGATTAAACGGATGGAAGTGCAGGACTCCAAGATT
TACTTCCTTCCCTTTCCAGCAGAAATTACCTGAGACGAGATCTGCATTTACTGCTCAACCA
CATCTAATTTGATGTCTCTGCAGATTTAAAAATGTGTCCTTCTTTTCCGTCACCAAGTC
ATCCCTGGGTACTACTGAACATCCTTCTCAATTCCCCCGACCCCATGGATGGCTGTTCCT
CCATTGTCTGTTTACCCAGATGTCTCAAAAAACAACAGACAGAAAGGAAGTGGCTAAT
GGAGCTGTGGAGTCCAAAGTGTGACTGCCAAGAGGAAATCCAGCAAAAGCCAAAAAGCCCAAG
CATGTAGCCCTGCCCCGAAGCACGCCACACGCATGGAAAAACCCAGAGGAAATGAGTGAGGA

TCAATGGGAAGAAGAGAGCCAGCCAGGAAGTTGAAGATTTGTCTCAGGAGCAGATAGCTGA
 AGAGAGAGAGAGAGAGAGAACGGCTTACAGCTCAGGTCCTCTCTCCATGCTTAGGAA
 CCACTACAAATGCTACTGCCCTTGAGTCTCATTTTGTTCCTCTGGAAACCATGTGTA
 CCTTGTTTGCAACAGTATGGGCTCACAGGCAGAAAGGAATTTTCCTTGCTTTGGATGAGAC
 TTTTGACTTGGACTTTTGGGTTAAGTTCTGGAGACCAGAAAGGCCAAATCAAAAGTATGG
 GCAGGCTTGATTTCTTTAGAAGACTCCAGCGGAGAACTGTGTCTCCTTGCTTCTGATTCT
 ACATCTCCATCCATGGGCCACTGTTTCAGCAACCTCAGCCAGTGCAACACACCTCAGCC
 AAGAAGAGTATGCAGAGAAAGGAGTCCCCCTACCCCTGCACAAAACCTGTGTCTGAAAACTG
 TCTCATATTGTCTCAAGTTGTCAATTCATTGTGAATTAGACCTGTTTAACATGTAATCTGC
 AACATGCTTCACTGTCTAATTTTCCAGAGCCCCCTCATATAAGGAACTGTATTATTGGTAT
 AATCATCATGGTGAAGAAGTTGGTATGTGGGGGAGAGATGACAGAAACAGAGAGTAAGTC
 AGAGCTGGCTGCCTGA

SEQ.ID.NO. 440 HG1015011N0 PLT00014333A08.contig
 AGGAGGATTAAGATGGCCACCAACAGCTGCGGGAACGGCAACAACCCCTCACTTTCCG
 GGATGGTCCCTGCGGGTCGGCCCGGCTTGATGGAGAGAAGAAACCCGAGGAGCGCCGAG
 GCTAGGCGGCGGCGGGACCCAGCGAGGACGAGACGCGGCGGAGCAGGGACGGGG
 GCAGGAGAAGGGAAAGGCGGGCGGCTGCTGCCCTGCTGCCTAGCACCGCTGCCTGGCC
 CGGCGGACCGGTTCCCATACCTCGGGCCGCAGAAATCGAGCTCGGGCCCCCGCCCCGGC
 CCGCGGCGGGGCTCCCGGGCCCCCGCGGACGTCGCGCGGTCGCCCTTCCCCCGTA
 GCCCGTGCGCCCTCGGCGCGGAGCCCCCGGCCCGCGGTCCCGTCTCCTGGGCCCTGTCC
 CGCCCGCGGCCCTCCGCGGCCCCCTCAGGTGAGTACCCCCGCCCATCCTCCCGCCCCCGCG

CTCCCGGAGAGGGGTGCTGCTCGGCCCGCGCGCCTGGCCGCTTGTTCGTGCGGGGC
GGAGGCGCTGCCCACTCGGGCCCTGCGCCCGCTCCCCCTAGGACACAGCCCGCGCCACAG
CTGCTGTCTCCCCAGGGTCCGCCCTTGGCCGGGAGCCGCCCGCCAGGGTGGGCGG
CCTGGCCCTAGTCCTCGCTGCCCTCCTGCGGAGGGCGCCGTCGCGGACGGGTCGGGGTCG
GGGTACGGGCGGAGCGGAGCGGAGCGGAACCGGGATATTTCCGGGGGATAGGGGAGGA
GAGGCTGGACGTTACCTGTGGAGCAGTGGGGCGTAGCGGCCCTGGCTCGCAGGGG
TAGGGGATCAGGCCCTGGCCCCCCCCCGCTGCATCCTGCGGGGTCTCAGTCCCGGAGCC
ACCCGGACCTGAGCGCAGCCGAGCGGGCCTCCATCCCTCGGGCGCTGTCTCCAGGGCTT
CGGGCCTCTCCCCCTACGCCCTACGCTGGAGGCCCTTCGATTATGGAAGGAGCCAGAGAT
GCTCCAGTTCTTGTACTGGGGATGTGGTGGCATCGCAAAACATTCACAGTTACGGCGGG
ATTTAAGAGCAGATGAGACCTTTCCAAACCAAGGCCATCTGGCTAGTGTTTGAGCCCTCA
GGTGCAGTTTTTTCATTGATTTTTACCATTAAATTGTGAGATCTGCAATATCAGGTCCTCC
TTTCACAGAAACAAATGTGATACAAACCAGATTGAAAGGATATTTACTGCATATGTGAACCC
TTGCATGGACTTCCAGTACCTTAAAAATGGTTAAGTGTTTTCTGTGGAAAATGGTGCTC
ATGCATACGAAAAACTATTTTTGTTCAGTAATGGTTGCCAACTGGTTTACTATGTAGTTT
GAACACCAACAATTAAGGAAAAACAACCCCTATTCCTCGGGCTTTTTAAATAGTG
GCATATTTCCAGGTGGTATTTGGTAGCAGGTTGAGAACAACTGGTTTCGACTATCAGGC
GACGTAAACAACATAACATAAAACAGAACTAAAAATGTGCTAATTTATAGAAAAACATGT
TGTTATTAGTGCATAAGTCTAGTAAAAAAATCCGACTTTAAAAAAATAAAACTTTATCAT
GAGACTCTTAAGTATAGGAATTAGGTTGGTATGTTTTTATGTTTTAAAAATATATTGCA
TTAGAATTGAGAACTCTTCCTATGATATTTGAGCAATAGTAAAGTATGTTAAATACTTTGA
ACCATCTATCTCAAGCTGCTAGAAAGAGGTTTCTACTGCTAGTTTCTTTCCATGAGTTA

AAATGGATTCACTGAAATATATGTATGCTTTTGGTGGGTACAGTGGGTGTCATGCCTGTC
 ATCCAGCACTTTGGGAGGCCAAGGCTGGGGATCACTTGAGGCCAGGAGTTAGAGACCA
 GCCTGGGCAACAAA

SEQ.ID.NO. 441 HG1015012N0 PLT0001433A15.contig
 ACAGCTCTGATTGGATTTCGAGGGATTAGCAAGAAAGGAGTGGCAGACGAAAGCTGGC
 TCCCAAATTCATGGCTGGCCTCCCTTAGCAAATCCCCTGCATTTGGTCTGTGACACCAATG
 AGGAGCAGGTCCCCTGTTTGTCTCTGCTGGCTGGCTGTCTGTGGAACAGTCTTCTAGGG
 TCACAGTGCACTGTGCCTCTGGGCAGCTGCCTGACCCAGGCCCTCCACGAGAGCAGGGAA
 GAGGCACAGAGACGTTCCATGATGGCTGGGTGAAGAACAGCAGGCAGTCGGGGTGAGTTAGG
 TTAAACCCAAAGCCCCCTTCTGCCCGGCCCTTTCCCTGGCTGTGTGCTGGGCTGGAC
 TCCCCCTCTCTGAAGCTCAGTTTCCCTGTCTCTAAAGTGGGTGTCACAATGACAGCCGCCA
 CCATTCCCTCCACTACCAACAAGAAATCATAGTGGTGACCTGCCCTCATATTCCCGAAACGTC
 TCTGCGGGCAACATCCTCCTTTGAGCCTCGTCTCATCCCATCGTTGCCCTCCTCCTGAAG
 GTCCTTCTGTCTGACACATCAAGTGTCATGGGACATCCGTAGCACGTCGCCAGCCAGG
 AGGACACTGAGGCCCTGCCCTGAGGCCCTATCTGTGGGTGTGATCATAAAGCCACCAAGAAC
 AGCTCTTGTGCTTACAGGGTGCTAAGTCTGACCATCTCAGGACATCCTTCACACTGCCC
 TACAAGGTAGGGACTCTCGCTGCAGATGAGGGAGGTATCTCTCAGTGGGGTAAGGGTTT
 GGCCAAAGCGGTACAGGTCGCCCTCAGGCTGGCCAGCGCGTAGCACTCCACAGCCACAGC
 AGACAGCATCACAATGAATACTCAGGAGCAGAGGGCAGGCTCAGTCCCAGGAAGAGG
 AGTGTGGCCACAAGCAAGATTACCCGCCCTGGCTGGAAGGGACAACCGAGCAGAAATTGA
 AATAAGCAGCAGGGAAGAGATTCTGTGTTTCTGTGGGGCCCTCATCAGAAACCCCATAGGG

CCCTTACCAGAAACCCTCTGCCTGCAGAGACCACCTGCAGATGGGGGCCTCAGCTGCCT
GCAGTTGCCAGGGCATGGAAAGCACAAGCTCATTTCTGTACCACATCACTAACTGTAG
GATTGGACCTCAACATGCATACTGGTTCTGGCACAAGAGAGCTGATTTCTTGTCTTGC
AACTGTCCAGGCAGGGTTCTCTGCCTCAGCAGATGGTGCTCTCTCTAGAAAGTAATTCAGG
GACCAGCCACCTTCCATCCTAGGCCTCCATCACAGCCCTTGGGCCCTCGCTGTGGTGTGCA
CCCCAGCAGTCAGAAAGAGGGAGACTGGGAGCACATCTCTGTGGCTGCAGGCGTCACT
TTCTGCTGTCTCTGTCTACACTTTCTGTGTTTTGTGGACAAATCCAGTCGTCAGGCCG
TGCCACCTGTCAATGTTGGATGGAAATGTAGTCCAGCAGTATGTCCAGAGAGAAGGAG
AAATGATTTTGGTAACAACCAACCATCTCCACTGCAGGTGATGCAGGCAACCCCAAAA
CGGGGCTCAGCCTGAGTTGGTTCTTGGCTTCATTCAGGGAAGAATTCAAGAGCAAGCCA
ACAGTGAAAGAAAGCAAGTTTATATAGCGTAGTGAAACAGCAAGATGGCTGCTTCATTGGC
AGAGCAGGGCTATCCACAGGCAGAGGGGCCAGAGTAGCACTTGTGGGTGCTGGCTGG
CTCTATTTATACCCACTCTTAATTATATGCTAAATAAGGGGTGGGTTATTCACAAACTTT
CTGGAAGAGGGGTAGACTTTATCCAGTCTATCACTGATGGGCACITGGGTGGTTCCAAG
TCTTTGCTATTGTAAATAGTGCTGCAAT

SEQ.ID.NO. 442 HG1015014N0 PLT0001433A16.contig
AGTATCCTCTCTGCTGGGGGTATTTGTGCACAGTGACAAAGTGTGTGTTTGC AAAAGCT
GATTGACTACTGTCAACCTCTTAGAAGGAGGTAAACCCATCATAAAGATTGATTGATGGCA
GGGTCAGACTGACTATGGCTTACAGTTGGAGCTCTGTTTGAGTTAAGGTCTTGCTCCCTA
ACTCAAAGAGAAAGTGTAGCCCTGGACAGGTAGCAATTAGCATCATCTGGAAGTTTGTGAAA
AAGCAGACTCAGGCCCTCATGCCAAATCCTACTGGATCAAACTCCATTTTGACAAGATTCCC

AGGCAATTACATGCCCCATTAAAGTTTAAGAGGCCCTGGGTAGGGTGCTTCITTTTGAAA
AGAACATAAGCATCTCTTCCAGAGGCCAGTCATTTACATTGTAATAAGATAGCTTAATCA
GAATACTTCTCATTTATCCCTTTAGCTAAAATATGCCAGAAATGGGAATTTTGTGGCAAGG
GAAAGAACATCACCAATCCCCAACTCCTGACCACTACTGTAACTTCAGCTCTTACCATG
GGGATCTAGCACATAGTAGCATCAACATATCTTTTGTCAITTTGGCTGAATAAACTCATTTCA
CAATGAGAAAGAATATGTGTTTTCTGTTATTCAGTATTTGACATATGCTCAGGTGTGGA
GGGCAGAAATAACAGACTACAAATGTCCATGTCCTAATCCCTAGAACCTATGAATATGTTA
CCTTACACAGAAAAGGGAGTGGTTTTGCAGATAAAATTAAAGTGGCTAAGCAGCTGATT
TTAAGATAGAGAGAAATGTCCTGGGTGATCCAAAGTGGGCCCATATAATCACAAAGTGTCCCT
GAAAAGTGAAAGAGGGAGGCAGAAAGAGAAATCAGAGTCAGAGAAAGGAGATGTGAGAAAGG
ACAAAGTGCTGCCCTTGCTGGCTTTGAAGACAGAGGAGGGCCCATGGTTGAGAAATGT
GGGTGATCTCTAGAAAGCTGGAAACAGGCAAGGAAACTCTCCTAGAGCTCTCAGAAAAAGAAC
GCAGTGCTACCAACACTTTGATTTTTTGCCCAATAAGACATGTGGTAGACTTCTTACCTAC
AGAACTGTAGGGTAATAAATTTATGTTGTTTTAAGCTATTAAAGTGTATAGTAATTTGTTA
TGGCAACAATAGAAAACCTGTAAAACAGGAAACAGTTCCCTTTGTTTGGAAATATGAAAAAC
AGACAAAATGCTAGAAATTATGTATAATTTTTTTTCTTTTTCATATCTTTTGGAAATTCACCTTG
CACATGTTTGAGGTAATCCTGCTTCTCAGTTCGTGGAATAATTATCATTTATTCATAGAAAG
AACCAACCTTTTTGTGAAAACAATTTTTTTAAGGTTTAATTGTTTTTTTCCCTTCATCTATC
ATTCCCTAATCCCACTGCCCTTGCACAGGCATCCACTATAATACATTTAATATAGTTC
TTCCAAATCCATCCAACAGACGTTTATTTAGATATATGAGCGTCCTTACAAAAATACACAA
GGCCGCTTTATGTGTGTTTGATACATAAAATGGAATTGTGCTATAAATCCCATTCGTGT
TTTGCTTTTTTACTCTTAACCTCAATTTTGTGAGATCAATCCAAGTTTCTAAATGAAAAATAT

AATTCATTATCCTATTGTATACACAAACTATAATTTACCTGTACATTCCCTTGGACATAGA
CACCTAGGTTGCCCTCTAGCAGAGTGATAGTGAAGCTTGTTAGGAACAATCTTGTGTCCCT
GGGAAGAATTTCCCTGGGTGTATACCCAGGAGTGGGCTTGGTGATTTAAGGTGCACAAT
TATCCTTCAGAACGGCTGCACTCTGCAGTGGGTGAGGGCTCTCTTTCCCCACCTTTTGGCC
AATGCTTGTAATATCTGACTTAACAAAGTGATATGCAATGGGCAACAAGTAGCTGTGAAA
GGAAAGTGCTATATAGATGACTTAGTGGAACCATGTCTGTAAATGCAGAATGAGCAGGAA
ATGTTAAAAATAATTCTTGGAGGAAGAGTCTTACCAGCAGGACCGATGTGGCTGGGAAG
AGGAAAGCCCTCCAGTCAAGAAGTTGTGTGTGGCTCACCAATTGAGTCATCTTTATCA
CCAGTGAGATCACCCCTTCTTGACTACAGCTATCTGGAACTCAGTCTAGTCTATTTCCA
GCCTGAGCTTTTGTCTAACCTTCCCTGGACATGTGAGGGACTCTCTTGACTCAGCTTCCA
ACTCATCTCTGCTTTAACTTCCCTCTTTGCCACTTTACAAATTTGTGTGATTTAATGACT
TTTGACAGACACATGCCCCCTGGGAAACCACGATTATTAACACATGAAACATTTCCATCAT
TCCGAAAAGTTCCTAGCATAATGTTTGCAGTCCATCCTGCCCTCCTGCTCTTAGCATCATA
GAACTACTGATTTGCTTCTGTCACTTCAGGGAAGTTCACATTTTCCAGAAATTTTATATA
AATAGAAATTGTTTAGTACTCTTTTGTGTCTGGCTTTTTCACCTCAGCATAAATTAT
TTTGAAATTTATTCAATTTATTGACATATCAATAAATTAATTTGTTTTTACTGCTGAATAG
TATTCTATTGTATGAATATACGGCAATTTGTCTTTCTTTCTTGAGGGACATTTGTTTGGG
CTATTATTATTAAGCTTCTATGGACATTTCATATATAAGCTTTTGTGTGGACATATGTTT
TAATTTTCTTGGATAAATACTTAGGAGTCCAATGGCTGGGTCAATGACAGGTCTATGTT
TAACATTTTAAGAAACTGCTGGCCAGGTGCGGTGGCTCAGCCTGTAAATCCCAGCACTTT
GGGAGGCTGAGACAGGTGGGTACCTGAGGTCAGGAGTTCAAGACCCCTGGCCAATAT
GGCAAAACCCCATCTCTAATAGAAATACAAAAGTTATTACGGTGTGTGGCACATGCCTG

TAATCCCAGCTACTTGGAGGCTGAGGAAGAACTGCTTGAACCTGGAGGCAGAGGT
TGCAGTGAGCCGAGATCGCACCACTGAACTCCAGCCTGGGT

SEQ.ID.NO. 443 HG1015016N0 PLT00014333B03.contig
GACAGAACCGGAAGAGACCGGGGGCGAAGGCGACAGGGGTCTGTGGAAGAGACCTGTCCG
CGGAGAGCGGTCCACGTTTTCCTGGAGAAAGACGAGGCCCAAGGCAGGAGCGCGGGCTG
CGCTGGGCCCTTACTTCGCCGCCGGGGGGGAGACCGGCCCGTACCCGAGGGGACG
AGGGACGAGGGCCCATGCCCAGTCAGGGAAGCCGAAGGCCTGGAGGGCTTCCGGGAG
CAGGGGCTGGAGTTCCTCTGCCAGGAGGCTGGCACCAAGACACCCGGCAGAGGGAGG
CGGCAGGGCCAGCAAGGATTCTCCCCAGCCCCCTGTGCTGCTCCTGCGGCTTCTG
TGCGCGGACCGTGCTGCGGTGTGTAGGGAACGCTGCCCTCTCTGCTCGGGAACGTGGA
CCTTTCCCTCCTCCTCGCCTGGCTACTTCTGACGCAGGTCGTCAGGACTCCGCTTGGGT
GTACCCCGTGCAGGAAGCCTCCCTTAGTCAAGGGCCCTCGGCACCCGCCCATATGTGAC
TAGCAGCCCTCCTGTGTATTTCATCGCGCCCCCGTTGTTTATATCAGCTATTCCACTCAC
CACCCCTTCTGGGCAACCTTTTATCTCCGCTCTGAAATCCCATCGCTGAGGCTGGGACC
CCTCCTCAGTGCTTATCCCTGTTTCCCGCGCGAGTCTGGCGCCTGGCGTGTGGAAAGCG
CTCAGTAAACGTTTGTGAGCGAAGAAACGACGCAAAAGGTGATAGCACGACGCAGTTAG
GAGGCTATTGGCCCCGGCGGGGAAAGGGGGAAAGGTGCGGCTCGGGGTGGCAGGACCC
CAGAGGCAGGGGTGACTGCCGGGTGCTCTTGGGGCAAGGTGGCGTCAAGGCCCCAG
CAAGGTTGGGATAAAGTTTCTCTCCGAGGCACAGACTGCCGCCCTCGGGCTCAGCATTTA
CCTCTCCCTTCCCTTCCCTAACGCAACTTCAGCGGAGGGCTGAGGGCTCTACTGCGCAT
TACCTGGAAACTTATTTCACCTCCACGGCCCTAAGAGGTGGCCCTTATTATTAGCCACGG

CCAGGGTTAGGGTGAGGCCACTCCTCCAGGGCAGAAATTTCAGGGGTGCCCAACCCCCCGG
 CCTAAATCATTCAGCAATCAACATAAATTATAGTTCAACGCAATATACTTATCTTGTG
 GGCCGCAAGCCACCCAGCTGCCCAGGCAAGAGCCTGAAGGCACAAGCTGTTCCAGTACAG
 CAAAGAAAATAATTAGAAATAAGAAAAGTTTACTAGAGATAGGAAACGGATAGGATTATA
 TCTGACTATTATTAAATCAATTAGTTTGTAGCATCACTCTTTGTTCTATTACCATAAATGATC
 TCTGTTCTATTATGATTACCTTGGGGGAACACAGGCCACACAGAGTTAGGAGCTGAAGGG
 CCACAGTGAGAGGTGACCAGAAAGACGAGAGTGTGAGCCCTCATTCACGCCCAGAGAAAGGG
 CCGCTGGAGGGCTCCTTGGCCTAGCGGTAATGCCAGTGCCTGGGAAGGCCCTGGTTACTT
 AGCAGGCCCTTGGTCTAGCGGTGGCCCCCAGTGCCCTGGGAAGGCACCCGTTACTTAGCAGAC
 CCGGAAAGGGAAATCTCCCTCTCTCCAGGGGAGACAGAGAACGCTCCGCTCCACCACTCT
 TGTGGAGGTCTGACATTAGCCAGGCCGCCCGCAGTCAATCCGGAGGCTCCAAACGTCTGT
 CTCCCTGTGATGTGTGCTTCAGTGGTCAACGCTCCTTGTTCACCTTTCATGTTTCAGCCTGT
 GCACCTGGCTCCTCCTTTTAAAGTTCTTAGAAGACAGCAGTAGCAGAACTAGTAGGAGTAC
 CACAGTCTTCGATCTTTCTGATAAGTGCATAGAAAGAACGCTGACGTTTGCTGTCTCTCCC
 TCTCCACCTCGGCTACCACAAAGGGAAGGCCCCCTGTCCAGTGGACACGTGACTCGCGT
 GACCTATCGATCATTTGGAGATGACTGGCACTCCTTACCCTGCCCCCTTGCCCTTGACTACA
 ATAAATAGCAGCGCCTCCAGGCACCTCGGGGCCACTACCTGTCTGTCTCCGGCCTTTGGTG
 GCA

SEQ.ID.NO. 444 HG1015018N0 PLT0001433B05.contig
 AAAGCCCAAGGGCTCTTCAGTTAGCAGATGATGAGTGCTTCCAGGACTGGGTTCTTCTCT
 TCAAGGCAACAGGTTTCTTCTGTGGCCACGGGTATGTCTAGAAATGTTGTTTGGATCTAG

GGCCTATAATGGGGTCCTCATGAGTCTGTCTGGTCCCCCTATTATATTGTGGTTGAACTAG
TGTCTAAGATGCAAGACAAAAGTTCTTTTTACTCTTTTCCTGTTCCTCTCCTCAAGCAGAAAAG
AGGGAGTCACTTTTGTCTGTCAAAGATGCACTACATGGGGTTGGGGTATGGGTGGTGCAA
GTACTTCCTTAGTTACCTTATCTGGTGTCTCATTTTGGTGTCTCTCCAAAGTCCACTTGACTCC
AAGCCCAGCACAGCACTAGGACTTGCCCAGACATTGCAGTCCCTTGTGGCCTAGACTGCCCCT
TTCAAATTTTATTTAGGACTCTAGAGTACTTTAGCCCCATGGTGGCAAGGCTTGTGTGAAACT
CAAGTTTCAACCCCTAGGATGGACAAATTCTCCTCTCCTCCCTAGGGCTAGTCTAAATGTTCCCC
TCTGTGGCAGGCAATTGGCTGAGTTTAGCCCTGATTTTGTCTTCTGTCTGTGATAGGGAAGC
ACTGAATTCAAATGCACTGGCTCACAGGACTCTGTGTTCCTCCTCTTAAGCACTCAGAT
TCTCTTTCCACACCATGTGGCTGGTGCCAGGGAATCAGGGAGAGTGGGTTCCGGTGATTCA
AGACTGTCTTTCCCTACCCTCTTCAGTGCGTCTTTTCAGTAATAAAATTAACAAAACC
AGATAATTATGATTGCTCATCTCGTTTTTGTGATTCTTATGAAGGTGATTTTTTTTGTGTGTG
TGTATAGATGGTTATCAAAATTTAGCATTCCTGTGGGAGATAATGGACAAATCAGTGGAGGC
CTCTATTTCAGCCACCTTACTCTGCCCCCTAGATTGTTCCTTACAAAATACACACATGCTTAT
ATCATTTCTGATGTTTAAACCCCTTTAAAGTTCCCCCATTTGTGGTGACCTGTTCCCTCTTA
CATTATCAAACCTTATCTCATTACTTTCACTCTAACTCTAAACTTTAGCTACATTGTCCCTT
GTTTCAATTCCCTCAACCTTGCCATGCTTTCTGTGACTCAGAGCCTTTTCACATCCTATT
CCATTTACCTGTGCTCTTAACCCCTCTTTTTTCTTTGTAGATAAAAGATGTATACCTT
CTGTGGAGCTGATTTCAATAATTACTTTTTAAAGGCCCTTTTTATTGACTCATCAAACTA
GATTTGGTTCATTATATGCCCTGCTTGTCTCCTACACCTTGTGTGTGTGTGTGTGCATTCC
AGTTTTTCTCAGACAGTCCCTGGTTTACTCCTATTGTGCAGACATAAATACTACCAATAGCACCC
TCTTTTTTGATCTTAAAAATGTTGAGGTTTGGATTGCAAAATTATGTTTCCCTACTTGTAT

ATAAACCTTACTGTGCATACACAATTGCAGTTAAACACCTTACTTGTATAAATTGTC
 TATCCCATTTTAAATTTAGGCTTCATAGGAAAGACCTGTGTCTGTTTTTGTCTGAAGCA
 TCTTGATCCTTAGCATCATGCTAGTGTCTTTGTAGGTGCCAAAATAAATATGTGTGAAAC
 ACTATAAATAAATAATATGCCCTGAATTAATAACAGTCCTCCCTTGGTATCCATGAGG
 AGTTAGTTCAGGACCGACCTCCCAACATACACCAAATAATAAGGATGCACAAAGTTCCCT
 ATATAAGATGTCATAGTATTTGTCATGTAACTTATACACATCCTCTGATATACCTTAAATC
 ATCTCCTATAAAATTACTTATACTATCTAATAACAATGTAAATTCTATGGGAATAGTTATA
 CTGTGTTGGTTTTTCAGTCTCTATTACTTTTATTTAAATTATTTTATTTTGTAAAGTGATTT
 TTTATTTCAAACCTTTCCTTTTAGGTTCAGGGGATACATGTGGAGATTTTGTACATGGGTAT
 ATTATGTGACACTGACGTTTGGGGTAGGAAATAATCCCATAAACCCAGATATTAAGCATAGT
 ACCCAATATATAGTTTATCAACCCCTTGCCCTCCTTTTCCTCTAACATCTAGTATCCCCCA
 ATGTCTATTGTCCCCCATCTTTATGTCCATGTGTAGTCAGTGGTTAGCTCCCACTTATAAG
 TGACAAACATGCAGTATTTGATTCCGTTCTGTGTTAAATTTCTTAGGATAATGGCTCCA
 GCTGCATTTCATGTTCCAAAGAACATGATTTTCATTTTATTTTATGACTGTGTAGTATTCCA
 TTGTGTATTTGTACCATAATTTTCTTTATCCAGTTAACTGTGTGGCGTCTAGGTTGAT
 TCCA TGTCTTTGCTATTGTAAATAATGCTGTGTGAAAGATAATGAGTGCATATGTCTTTTT
 GGAAGAACAAATTTATTTCCCTTTGGTTATATACCTAGTAATGGGATTGCTGGGTCAAATG
 GTAGTTCTAAGTTCCTTTGAGAAATCTTCAAACCTGCTTTCCCATAGTGGCTGAGCTAATTTA
 CATTCCCATCAACAGGGTATATATTTTGGTTGAAATTCATGGATGGGGACCCAGGGGATAT
 GGAGGGCCGGCTATATCAAAATGACAAATGAGAAAGCAGCAGTTGAAGGAGACTATACTGGC
 ACAAGGGCAGAGTGGTACATAGTATTTTACCACGGTATTGACAGAAATAATTACAAGT
 TTTAACTTTGGCTTTAAATAGAAATCTATTTTACGAAGTTGATCTTTTGTGCTTATTATT

GCCATTTACATCTGTTTTTCTTTTTTATTACCTTAAATATACACAATTTTTATTGCCA
 TTTTACATCTGTTTTTCTTTTTTATTACCTTAAATATACACAATTTTTATTGTCAATTA
 TATCTTA

SEQ.ID.NO. 445 HG1015020N0 PLT00014333B15.contig
 CATAACACACATAACACAACCCATTAAATTTTTTTTCAAAAAGCGAAGTTATTTTACT
 TTTTCCCATTATGAAAGTCATCAATGCTCATTATAGAAAATTTGGCAAATAACAAAAAGT
 GGAAAGAAAAAATTCCTTAATAATTTAGTATATTTGATATATTCAATCCTTAATCCTTAA
 TCTTTTACTATGTTATATCTTTAATAGTTGTGATGATATACTTTTTTGTAATAAATATATA
 TTTGTGTTGGCTTTTTTAAAAATTTAAAAACATAACAAAATATTTCCCCAAAATTTCTATA
 TTGTTTGGGTAACAAATTTTAAATAGCTGCATAATTGTTTATATATAGCATAGTTTATTTA
 ACTGTTTACTCGCTGTTAAGTAATTTAGGTAGTTTTCATAATTTTTCTTTCTTTTTTTTT
 TTTTTTGTGATGGACCTGGTTGAGAGTGCAGAGACTCCGCTAAAAAAAACAAAAAGCC
 TTCTGTCTGT AAGTTTTCTGACTGCTGCTTTTGTCTCTTAATCACAAAAGTAATCTGTAG
 AAGAACTTCCAAATATTTGATCTTTCTGTTTGTGTTGAATCCGCATTGAACGGAACTTAGT
 GTAAATATGTACTTGGAATACTCACTACGTGCTCCAACAAAACAGATTGGAAAGATGCAGA
 GTGGATTGCTCTATCAGTGGCTCAATCAAAATGAAATGTTTACTTTCTGCCCTTTGAAAAA
 AAAACCTGTTTTTTTCTCTGGAGCATGATTCCCTTTCACACAAAAGCTACTGACATCCTCT
 CAAATTAAAAAGGGGTAGAGAAATACGACATTTTCTGTTTATTACTGTGAAAACGTGGC
 CTCAAATTCGTACCAAGATGGCCATGAGCTCTGTCTGCCATGCCCGTTTAGGAACACTG
 CTCTTTGGGCTGAGCTCATTAAACCTGAAAGACTCAGACAGACAGAAATACTTTTAGAAC
 ACACTGAAGAACAGTGTGAAATGATGTGTCTATAGCTGTGAATGGCTGGAAAAATCCTGAT

TTGCTAGCATTATAACTGCTGAAAAACAAAAAGAAATAAGGAAACAAAAACAACTGCTG
 AGATGGCTGTGCACATTGAAATAGGCAAGGAGAACTCTCTACCATATCTTTGATGACAAA
 AAAACAGTGGAGGAAATGCAGTGAGTCGATTATATTAATATTTCACTTCATCATAACTAT
 AATGACAGCTGACCCCTTACTGAGTACTTATTATATACTCAGTACACTCAGGCACTGTTCC
 AATCATTTTACACATATCAATAATGAGTAGCGACTGTTATTATCCCCATTTTACAGATAT
 GAAACTGAGGTATGACGTAATTACTTAACCTTGCTTGAGGTCACACAGGTTGTAATGGAAG
 AGCGAGGATTTATACCAGATAGTCTTGATCTAAAGCTTCTGTTCTTAACCATGGTGCTTT
 ACTGCTTCTCACAGTATTTGACACAGATAACTTAATAAGCTTGTAGGAATGTAGGAATACA
 ACCTTTGGGAAAGTCATCTTGCACACCTGTTAAATTTAAATTACACGTACCTTTATTT
 TTATGTATTTATTTGTTTATTTTGAGACAGTCTCACTCTCACCAGGGTGGAGTGCAGTG
 GTTGAGCTGAGATTGCCCGAGTGTACTCCAGCCTGGGCAATAGAGTGAGACT

SEQ.ID.NO. 446 HG1015021N0 PLT0001433B17.contig
 CAGGTTTCTTTTTTCTCTTCTTTCTTTATTATTATTATTTTTTGTCTATCACTGTACCTT
 CAGTGACTGGCTTGGAATAAACACTTGATAAATAAATAATATGTTGAGTGAATGTAATGATG
 TGGCCCCACCCCTAGTTCCTAAACCGAGAAAAAAACCCCAAAGCCAATTTCACTCAA
 AATGCTGATTGAATGACAGAAATCCATTCTGTGTTTCATTGGCATGCTATTTTCAGAAAGAAA
 GAATGGAATGTATGAAGCCGTTGCAAGGTGCAGCATGGATTTCATCTTCTTCTTGAAGCCT
 TCTTCCAAAGCACAGTTCCTGTTCTGAAATTAATGCTGTTTGATGATATGGATGATGC
 ATGCTCTAGGAATTTTACCACCATCACCATCTCCGATACAAGGGATGCATAGATAGATAA
 TTGATGTTGTCATTTAAGGGTTACCCACGAATCTTTGGAAGCCATGATGAGGATATGCCA
 TTGGCTACCCCTTTTGATCTGACTCACTGGGCCCTCAACAGTGCCCCCTTCTTCTCAACC

TTCAATCCAGTCCCTTCTGCCTGCTATTTCGGAGTTTGTATATAACCTCTTTCTTTTCATGCCT
 TGGGTTTGGAGACGTTCTGCCCTTCATTCCTTTTGTGGATTCTTTGAGATGTCCAGATGTCT
 TCAATATTCTTTTCTTTCTCTCACCATAACACATCCATAGCTTCACATGGTCAG
 CGTGACCACTGTGAATAAAAGACTTGACTTTTAGCTTCTCTTTTGCATAGATGAAGTCA
 GATGCTGTGTCCGTTCAATGGAAATATGACTGAAGATGATGCTTGCAAATTTTCAAAAT
 GGACCATGACAAGGGGTGAGGTCTCAGGACCTCCTCCCTCCATGCTGTTAGAAAA
 TGAGCTGTCCTGGACTGTGTGGATGAAGGAGGACAAAAACAGTTATGATAGTGAAGCAAAA
 ACACAGGATGCCCCAAGCCCTGTTGCCATTGAGCTCTGGGCTGCTTGCAATCCCGACAGCCT
 AAGGAAAAACAGACGGCTGTCTCATTTAAGCCTCTGTGATTTCTGGTCTTTTGTGCAGCA
 GATGCAACTTTCTTCTGAGCGATCTTCTCACCCTTCATTTACACAGCCAGTGAAAAAGAA
 TGGAAGCACAAATTCTGCGTCCATTCAATTGGACATGTTGCCAGATTCTTACAAGATCCAT
 GAGAAACAGGGGCAGAAAGTGAAATAAACGTTTGGAAAGCATTTGCCACTTAACAGTATACC
 TGTGGGAGCAGCAATATGAAAAGTGGTAATTGGAAGGAGGAAGCGCCTTGCAATGGTGAT
 TGTATTAGAAAGGACTTGGTGGAGGCAAGCCTGCTGCAGCGATCTGTTCAGCCGTGGCG
 GTGTCAGAAATGTGGCAGTTCTGACGTGGTGGCCCTTTGTGTCTCAGGAGCAAAGTGACAA
 AATGGCAAGGATGGGGGTGCTCATCTCAACCTCACAAATGAGACGGCACCTGGAGAAAC
 ACATTTTCATCCCAGACACCTCATTAGCAGAGAGACACAGATAAATTGGACAGGGTTTGGGA
 GAGGAATGAAAAAAATTCAAAATATAAAAGAGCAACACTTTAACCAAGGTACTGAAGGA
 ATTGATTTATGAAGAAAGATTAAAGGAACTAAATATGTACTCTCTGGCTAAGTGGTAACTA
 AAAGAGAACATAATAGAAACATATTTGAGAGGGGTAGAAAAACAAGATACATGAAAAAGGGATGG
 TTTAATTTAGTCTCAGAGGGGAGTGAGTAAAAACAATGAAATTCAGAAAAAGAGAAATGTA
 CTATAAAGTCCAGGGAACAATCCAGGAGATGAATCTATTTTGATCACACTAATCTTGCA

GCAGGAGATTCACTACTTGATGGAATTAATCCTACTGGTAAGAACAGAAATAAATGTTTT
 ATGTTGAGCTGCGATAACTCTGTCTGCTCGCCGAGGACTGTGGCTGTAGATTTGCTCAA
 CCACCCATGACGATAAGCTTTGGTCCACCTGCCCCACAAGTCATCAAGCAGCCAAACCAAAG
 GAGATCAGCTAGTTGTCCTGGAGAAATGTATATTACCTGCATCACCTGTTCTTTCTCTAGGG
 TATTAGTTGCTGTTGAAATATGTGTGGGGTCTACACGGGGTAGGAGCAGGAGATAGCA
 AGAAGGACAGCTGGAAGTGAATACTTATTGTATACCTTGAATCGTTTGTCTTGCATATG
 CGGGTAGGTACCAATTGTCCATCCCTTCCCCCATGTATTAGTCTGTTTTCATGCTGCTGA
 TAAAGACATAACCGAGACTGGGTAATTGTAAAGAAAAAGAGGTTTAATGGACCCACAGT
 TCCACGTGACTGGGAGGCCCTCACAAATCATGTAGGAGGTGAAAGGCATGTCTTACATGG
 TAGCAGACAAACGGAGAAATGAGAGCCAACTGAAAGGGGAAGCCCCTCATAAAACCATCAGA
 TCTCATGAGACTTATTCACTACAAATGAGAACAGTATGGGTAAACTACCCCATGATTCC
 ATTATCTCCACTGGGTCCCTCTACAAACATGTGGGAATTATGGGAGCTACAAATTCAGAT
 GAGATTTGAGTGGGACACAGCTAAACCGTATCACCTCACAAAATAAAGAGCTTCTGACT
 GATGTCCTCTACTTGTTAAAAAAGGAAAGAAAAGATGAGTGCTTCCAAGCCATTCTGA
 ACCACTGCAAGAAAGGTTTCTTCTAATTTAAACCCCGAGATTGTGTATCACAAGGCACGGCT
 GAAGGAAGGCATTCTGATAATGCTTTCAGGAATGTACCAATGAGTAATTTTGTGTCTAC
 TTTCCATTTTGGGGTATGATTCCAGAAAGCAAATAATTAATATATTAAATAAACATTAT
 TTGGATTTTATAAATGAGTATAACACGGTGAAA ;

SEQ.ID.NO. 447 HG1015023N0 PLT0001433C02.contig
 ATGCAATGCACCTTCTTCAATATAGAAATTTCAATTTTCTATTTTCATACAGTTTTCCTGAA
 CTATAAATTTGTGATAGTTGTTCTATTCCTAACTTTGTATATTTTTTCCAGAAACTCCT

GTAGCCATATTTATATCCTCTTTGTCTATCTTCAATATTTATCAATTTCTCTTTGTTT
TTTCATATAGTTCTTCAATTTTTAACCTTTAAAAATGTTTTTCCCTTTCCCATCTTCTATTT
CTTTAAGGAGTTTGTGCTTGCATTTTTTCTTGTTTAGTTTTAAGAAATTAATTTATTTTC
TTTCTGGCTCTTTCCCTTAGTTCTATGACCACATCTAAAGTTTTTAAAAATCTGATTTGT
GTTATTTTTTCAACTTTTTGTACAAATTTTCTTATTGCCITTTTAGCTCATTTGTGAAAGAGA
AGATTACAAATTTTGAICTTCTCTATGGACATGGCTTTCTGCCATGCTTTTCATTTGTCTGTA
GGGATATTATCTTGTCTCTTTGTTATCTTTTTTCTAAATAGCTTTGTGTAGGAAATGGTTGC
TATCAATTTCCCACACTCATTTTATGTGAAATTTGGTTTGTCTGAACTTTTAGAAGGAGGCA
TGGTTCAGGGTAGCTTTTTTAACTTCATACGACCCCTCACTTCTGTACCTGAAACAATAA
AATGTAATGAATACAGTGTTCAAAAATCAAGTGGCTTGTCTTCTGAGATTTTCTTGGGTGT
GTCCCTCTCACAATGTAATTTAGATCTTCTCTTTTATTATCTCTCTTACTACCAGCACCTT
TTCTTCAAAGTGGAGTGTCTCTTGGAAAAAGGCCCTGGTAGACCAAATTTTGAGAGTTCA
TAGGTCTTTGGGCCTCCTGGGTGTGAGTCAATAGCCATATTCCCAGGTGTAGACTATGTTG
CCTCTAGAAATTCATAGAGGCCCCACCAGGCGCTGTGCTTCCITTTGTGTAGGGGATGC
AAGATGCAGCAATATGCCCTTACTTTGGAGGATATGTACCAACAATCCTGTCCACTAG
ACCCCTAAAAACCTTCTGTATTTGGCCGATCTCTGGATCTTTGTAGGATTTATTTCCCAT
TCTCTGGAGTACATGGGTCTCACCAAGGCTTCCAACTTGCACAAAGCCACCTCTTACTCAT
CCTTCTGACCAGTATGATATCACAGATGTAATGAATCAATGTGATGTTCTACAGCACTG
ATTACTAGACAAAAAATTAGAAAAATAAAATTTCTAATAAAAAATTAGAAATAACCTAAAT
ACCCATCAATGCCTATGACTGAGTAAATTATTACACAATATTAATTATTATACAATAATT
TACTCAATCATAGGCATTGATGGATATTTAGCATATTTCATCAATGGATATTCAGCATATC
TATATTTAGGCAATTGATGATATTTAATATATCATAGAAACCTCTGTGAAATCTATAGATGT

TTCTTTTTTTTTTTTTTTTGGATGAAGTCTCACTGCCATTGCACCTCCAGCCTGG
GTGACAGAGTGAGACTCTGTCTGAAAAAATAAAGAGTGAAAGAAATAT
GGCGGTAAAAACTGGAGAAAGAACTTTCAGTTTTAAATTCACAAAGACTGTCAGAAATT
CTGTTGCTCAATAGTGCCTTGGGATAATATGATTTTGAGGAACGTGACCATGGTTTTAG
CTCACATGTCACATATCTGAAACAAATGAGCCCCCTGAAATTAAGGGGAAAAATGAGCATTGA
GGTCACATACACTTTTCCACAGAAACATGGCTTCAAGATTATTCATTTTACTGCTACTTAG
TATTTTCATGAAGAGGTGTATCTGTTTATTTAAGTTTTACCTGCCCTCTTGAACATGAAAG
TTACCTGCAATTTCCCTGCTTTCAATATCTCTGCTATAAAATATATTTGAATATGTAGGT
TTAAAAAAATCTTTGGGATTATTCCCTTAGATAAATCCCTCCCTGGGTCCAAATCTGTGAT
CCTCTCACAATTTCATGTTACCTATTATCCTCCAGAAATGACTGTCACTCCCAAGTGAAATT
GCCAAAAGCAGTTCCACATTTCTCAACTTCTCGAGTCCTATTTATTCGTGTTAAGTTGA
ATTTATTTCTTTCACATTAAGGCCAGAGAGTTCTGCCTAGACTACTTAGAAATGGTCAGT
GTGCCACTTAACACCATTTGGCTTTCCCTAGCTCTCAGACCTTCAAACTCAAACTGGATCTA
AACCACCAGCTCTCTGGATCTCCAGCTTGCTGACTACAGATCCTGGGACTTCTTGGTCT
CTGTAAATTGCTGAGCCAAATTCCTTATATTAACATATATCTCCCATTTGGCTCTGTTTCTC
TGGAGAAATCCCAGCTGATACAATAAGATAGAAAACTAGATAAATAAAGATAATTATAATTA
AAATAAATATTTTCAGTTTAATTATTAACATATGATATGCATAGGTATCTCTGTAAACCGATGC
TCTCTGAATAAAATATTCAATAAAACATGTTACAGTGTGCATATAAACACATGGGTGCAT
GTAAAAAATAACTGAAACTGAGGCTGGCGTGGTGGCTCACACCTGGAATCCCAATACTT
TCGGAGGCTGAGGCAGGCAGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACA
TGGTAAAAAACCCATCTCTACTAAAAATACAAAACTTAGCCAGGCATGCTGGCACACGCCT
GTAAATCCCAGCTACTCGGAGGCTGAGGCAGAAATTGTTTGAACCTGGGAGGCAGAGGTTG

CAGTAAGCCGAGATCGCAGCGCTACATTCCAGCCTGCGCA

SEQ.ID.NO. 448 HG1015025N0 PLT00014333C10.contig
 CACCCGCACAGCAAGGAGGTGAAAGAGGTGAAAAAACTCCCTTCTGAAATTGAAATCAC
 AGACCTGTGACCTGCTCTCCTAAATGGGAAGGGGTGAGGAGCGCTGTGGGAAAGGC
 TCCTCCTTGCAAGGAAAGGGATGATTTCTCCGAGTTCAACGGACTCCCCCTTTCCCCA
 GCGCCAAGCATCCGGGTGAGTGTTTTGCAATGGGGTCTGTCACCATCAGGAACCGAGGAG
 AGTTTAGACTCTGCATTTACTACACGACCCACAAGCACCACTGTGCTCTCAGTCTC
 AGGAGAAACAGCCTCTTTCACITTTCAAAATCAATCTTCTGTGCATGTATTTACTATAATG
 TTACAGCACCATTTGGTTCCAGCTTTAAATCACGAAGATATGCTCTACTTAACGACTTTAG
 TTGAAGCAGCACACGCCACATGGGAAATACCAATTTAGAGGCCAATCATCTCTTTTCCCAG
 AAAATCTCTGGGACAGGGGCAGGGGAGGGAGAGAGGGAGCTGAGAAAGCCTTGATGTGCA
 AGTTCAATTTCTACAGGTGAACACCAAGCGGCCCTCGGCAATGCTAGGTTGTTAGGAATTGG
 GCCTTACTTCTTTTCATCTCTGTGGTTTACTAACTGGCTCTCAAAGGCGTATGACTGCGG
 TGCCCCACTGCACGCACGCAGCAGAGGGCATGACAGGGTTTCATTGGAGAAAGCTGGCTCCC
 GCAGGGCTGCTGGGCTGACCTCTTTTCAGCTGGAGAAAGCCATAAGCGCTCACTAACCTGC
 CGTGCTCACCATGAACCTCCACTTGACCAATAGGCATCCCCCTCGTGGAACCTGCCCGAT
 GCTTTGTTTGGGGAGCCTGCTATAGTCAAAATTCAGGATGTACCAAGACATCCACGGAGAC
 ACTGGTGATCTCAAACCTCCCTGCTGGTGCTCTCCACAGGCCACAGCCACAGCCCAAT
 GTTCACCCCTGTCCAGGATGGCACCTGCTGCTGTGGGGCACGGACACCATCCGTGTGCAC
 GTTGTTCGGCTTGACATCGGGCCCTGGAGTCTTCTCTATAGGGGAACATGAAGACATCGGTT
 CAGTTGCCTGTTTCCGCAGCTATTCTTGTCTCACACTCAACGCTGCTTCAGCTGTTTCATGG

GACAGCATAAAATACACCTTCCACACTCTGTCTCAGGATGTTTAACTGCAGTCAGGGAA
 AGGCTGCATTTTCAGAAAGACTCTGGCAGGTCTGAACACCCATTGGCCACACTGGTGCAT
 GGGGTGCCTGAAACAGGTGCAAAACCCGACTCTTCACCTCCCAGAGTCAGATTAGAAATTG
 GAGGCCAGTCTGCTGTTCTATTAGTGAATCCTGGGAAGGCCTCTTAGCTTTAAGGGGTGT
 GTGTGTGTGTGTATGTGTGTCCATTTTTTTTAAATTAATAAGTATTTTAGTTGACA
 CATTATTATGGATATGATTTGATGTGTGATAAATGCATCTGTTCATGGGAAAGGGGT
 GTGGGAAGGAGACGGGACAGGGAGGATAGGGAGATGTTGGCCAAACAGGTACAAAATGACA
 ATCAGAGAGAGGAATGAGCTCTGGTTTTCTCTGCAACAGTGGGTGACTGTTTCCAAT
 ACTGTATTGCATATTTCAAATAGCTAGAAAGAGAAGATTTGAAATGTTCTCACCCACAAAAGA
 AATGATACATCTGGGGTGACAGCCATGCGAACTATCCTGATTTGATCATTTTACATCAAG
 GCCACTTTTGTGGGACCCGAACCAAGCCCTGAGGTTTGAGACTTGGTCACTGGAATACC
 TTGTGCTGGGCAAGTTCCCCGGGGTTCT

SEQ.ID.NO. 449 HG1015027N0 PLT00014333C16.contig
 GCTGCCATGCAAGTTAGCTCTTGGCCTTCTGTCTCTTGCAAGGATCCCAGGTCAGAACAC
 AGGAAATACAAAGGTGTGTACAACACTCACAGCTGCTGTTCCGCAGCCGGCTCTATGGGGAC
 CTGGAGGCCATCCTGTACCACGTGCACCTCTTCCAGCCACGGAGCTGCTGTGCAGCAG
 GCGGTGTTCTTCTCGGAGACACTGAGCGGAGGCGGGTCTTCCAGGCTCTAGCCAGGATC
 CACGACATCTGCTATAACAGCACCAACCCCTCTGGGACGTGACGGTGAGGGACCTGCTGCCCC
 TCCTCTGCTATGATAAAGACTTGAGCCAAAGATTTGGGATGCCCTTTCGCAAGAAGAA
 CTCACAGATGAGAAACTGTTTGCCCTACCACCTCAGCCTGCCCCCCAATCTTGAGGACTAC
 CACAGTCGGAACCTCACCCCTCACCTTAGAGATCCACGCCCCACAGGAGAAAGTACCTGCAG

TGGCGAGTGCCATGCTTATGAAGAAACAAAGACAAAAAGCACAGTTTCATCCAGAAAAAAT
 ATCACAGAAAGCCTACCAAGTCAGCAAGAAAGCCTCCGAAGTCCGTGGCGAAGGTGATTAAA
 ATTTACGCCCTTGCCAAACAAGGCCGTCTACAACTATAGTACCCAGACCATGAATTCTACA
 GAGCTTGCCAAAGAGGAGCTGTATCAAGAGATAGCCAAAGGAGCCAAAGAAAGAGATTTCACG
 TACTCACAGGATTACCTCTCAGCCATGGTGGAGCCCCCTGGACTTGAAAGGAAGAGGAGAAAG
 AAAGCCCAGAAAGAAATCCCGCCAGGCCCTGGCTACAGCCAGGGGATTCCAAGTGACAGGT
 CTTCAGAGCGACACCGAAAGCAGCTTTTCAGGATCTCAAAGCTGCCACCCATCAAAGAGCTG
 AATGAGGAGTGGAAGAAACTCCCTGTTTGCTAATGTAAGTCTGAGCCTGTGTTGGATCGA
 GACAGGTGGAGCTGGGACAGGCACCAACGTGGACTTTTGATCTGTACAAGAAACCAACCCACT
 TTCCCTCGAGCTGCTCCCTTCGCCCGCACCAAGCCTGTAAACAGTCAGGAAGAAAGAAAGGG
 AACAGCCCCCATCTCTGAGCAGCACAGACCCCTCCACCGCCACCGCATGGTGAACCTGCA
 CAGCCTCCCCCACACCCGACCAACCCCTCCTCAACAATCAACTTCATTAAAGTGCAGCAG
 GACAGATGGCAGCAGCCAGGCCCTGTGTGAGGCTGGGCTGAGCCTCCTCGTGGTCTGTGG
 TTGCGGAGCCCCAATGCGGATGGAGCGGCTGGCCCCGCGACAGCACGGCCGTCATGCCATAC
 AGGTTGATGAGGATGTTGGCCACCCGCTTCAGTACCAGCTGCTCCTCCATGATGGTCTGG
 GATCACAGAGGCTCCAAAGTGGGGACTCACTACCTAGACCAGTCCCCCACAATGGTCCCTCC
 CTGGGCTGCATCTTTGCTGTCTTAGTCTCCTGTGTCTCCTTGAGAAAGTGGAGTCAATAA
 CACCT

SEQ.ID.NO. 450 HG1015029N0 PLT00014333C21.contig
 CCATCGCTGCTGGAATCAGCCGCGGCTGCCGAAAAGTCCTTTTCGGCCTGTGAGCC
 ATGGCCGAGGCCCGAGCGCCCGGCCCTAGAACCCGCTGAACCGCAAGCCGCGAGCTTCA

GTGCCCCCCCCTAGGGGCGGCTCCCGCGGCCTGGCCGCCGAGTGCGCGCGCCCCGTCC
CGCCTGCCCTTCTGCTCACCGCTACTCTGTCTGCGCCTGCGTGCTTGGCTCCCTGGATCT
GGTAGGCCGAGTTCGTGTTCTGGGGCCTGAGTTTGCAAGCATTCCTTTGCCAGGTTAGG
TTGCGTGGACTGACACCTTAGAGATTTTCACGGCCTATCCGGTCTGAGTCCCACGGAGAG
CAGAACTCACGTGCTGTCCTTGTGCAACCCCATCCAGAAATCAAGCGTTATGCTTTTCCG
TCCTGACTGGAGCCAGATCCGTTGTCTTTACTTTCTGGAGAGAGTGCCCCCTCCAACC
AAGATTTAAACAAACGAAACCCACGAGTGTGAGGCTCGGGGTAGGAGAGATGCCTTT
GAGAACAAAACCATAAAGTCTGTACATAAACAAATGCATACAGACTAGCAAAAATCAAATG
ACTTCCCCCAGGCAGCGGCTTCAAGAGAGCCTGTTAAATGTTTTTCAACCTTAAAGAA
ACTATCTAAACAAAAAGCTATTGTCTGACGCAAAATTTGTTCTATATTCATCAAGCTCT
GTCTACTACTGTATCTGTATGTTAGAAATACTATAATAATGTTGAGTACTGCACACCTGT
TATCAACTCCATGTTTAGTGGCGTCATATTGAGAGGTGACAGCGGGCCAGCTGGAGTTCC
GGGTGGGCATGGGCTTGCGGGCCTGCACTCGGAGCAGCCGGCCGCTGCCAGCCCCG
GGCAGTGAGGGACTTGGCACCCGGGCCAGCGGCTACAGAGGGTGTACTGGGTCCCCCTAGC
AGTGCCGGCCCACCGGCGCTGTGCTCGATTTCTGCGCGGCCTTAGCTGCCTTCCCCGCGG
TGCAGGGCTCGGGACCTGCAGCCCGCCATGCCCTGAGCCTCCCATCCATGGGCTCC
TGTGCTGCCGAGCATCCCCGACGAGCACCAACCCCTGTCTCCACGGCGCCCAGTCCCCATC
GACCACCAAGGGCTGAGGAGTACGAGCGCACGGCCCGACTGGCAGGACGCTCCACCT
GCAGCCCCGGTGCGGGATCCACTAGGTGAAGCCAGTTGGGCTCCTGAGTCTGGTGGGGAC
ATAGAGAGTCTTTATGTCTAGCTCAGGATTGTAATACACCAATCAGCACTCTGTATCT
GGCTCAAGGTTTGTAAACACACCAATCAGCACCCCTGTGTTTAGCTCAAGGTTTATGAGTG
CACCAATCGACACTCTGTATCTAGCTGCTCTGTGTTGGGCCCTTGGAACCTTTATGTCTA

GCTCAGGATTGTAAAAACACCAATCGGCACTCTGTATCTAGCTCAAGGTTTGTAAACCC
 ACCAATCAGCACCCCTGTGTTTAGCTCAAGGTTTGTGAGTGCACCAATAGACATTCTATCT
 AGCTGCTCTGGTGAGGCCCTTGGAGAACCTGTGTGTGGAACCTCTGTATCTAACTAATCTG
 ATGGGGAATGTGGAGAACTTTTGTATCTAGCTCAGGGAATTGTAAACGCACCAATCACGGCC
 CTGTCAAAACAGGCCGCTGGGCTCTACCAATCAGCAGGATGTGGGTGGGGCCAGATAAGA
 GAATAAAAGCAGGCTGCCCCAGCCAGCAGTGGCAACCCGCTCCGGTCCCTTCCACAGTG
 TGGAAGCTTTGTTCTTTCGCTCTTTGCAATAAATCTTGCTACTGCTGACTCTTTGGGTCC
 ACGTGGCTTTTATGAGCTGTAACTCACTCACCGTGAAGATCTGCAGCTTCAATTCCTGAGCCC
 AGCGAGACCACAAACCCACCAGAAAGGAAGAACTCCAAACACATCTGAACATCAGAGGG
 GCAGACCCAGACGCGCCACCTTAAGAGCTGTAAACGCTCACCGCGAGGTCGCGCAGCTTC
 ATTCTTGAAATCAGTGAGACCAAGAACCCACCAATTCGGGACACAAATATTGGCACCTTGA
 AATCAGCCATAAAGGGACTACTTATACCATATAAATTGGAAGACCTTAGAAGTCCAGGTG
 TTTCTCCAGATATCCAGTCAATTAACATTTACCATTGTATCACTGACTATATGTCACCA
 AAGGAATCGAAGAGGCCAAAAGTTTCTCTGATTGACCTAGCCCCGGAAATTTTTCCTT
 GGCTTTGTCTTTTTTCTGGCTTTTCTCAATAAACTGAAACTCCTTGATGTAAATTTGG
 GAAATTATATGATACAGTATTATCAGTTATTCACAAAGTTATTAAATGTCCTATTATAAAGA
 GAGGCTTCTCTTTTAAATCC

SEQ.ID.NO. 451 HG101503IN0 PLT0001433C24.contig
 TCTTCCATTTCCCTCATTCATGTTGGAAAATGCTTTTAAAAGTATTCGATGACCCGTAT
 GTGAGGAAAGAGGAGATAGGCCCTTTTTCACCAAAATGTGACTTTAACCTCCTTAGTAAT
 CCAAGAAATGTATCAAGTTTAAAGTCACCTTAACAAAAGATGTTCTTTTAAAGGTGCAATTTT

AAAAAATCACTTTTAAACTGGGTAAACTGGGTAAAAAGAAAGGGGAAGGAAAAACAATG
 TGCCAGGCATTTGTATTGACACACATCTTGCCTGAAGAGGTGTTAGATTCCAGGGTAT
 TCCCTTATATTCTGGTATAAATGTTTGAATCCCGTGTATTTTGTTCATAATTTTGTGGTT
 TATCTTTTATTCTGTCTCCCTTTTCTTCTTTTACCCGCCCTTGGCCCTTAATGCTGCAGAT
 AGCCTCTGTCATTCTGAATTATTTTTCAAATCTGTCTTCTCCCTAATCTTTTTCCTGG
 GCTCTTCAGCTGCTCAGGTTTGACAGTCAATGGTGTAGTATCTTGTCTATGCAGGAAGA
 AGAGGATACGTGAGATTACTTGCAATCAACTAAAAATTGGTCACCTTTCTTCTCAAGTTTC
 CTTCATAATACCCAGTTTTAGTTTTTTTGTTTTTTTGGGGTTTTTTTGTCTTTTTTTTT
 GAGACGGAGAGAAATCGCTTGTGCGAGTGAGCCAGATCTCGCCATTGCACTCCAGCCCGG
 GCAACAGAGCGACACTCTGTTCAAAAAAGCTATATATATTTTTTCTGGCT
 AAAAATTAAATCCAAAGTGAATAGGGGTGTGTGTATATAATCTATTGAATAAATTCAA
 ATAGATTTCTAGTCTGAACCTTTGTGTGTGTGTAGGTGTTGTGTACATATGTAAAGTAT
 ATGCATATATGTATATGTACACACAAAAATATGCATATGTGTGCACACATATGTGCGCAT
 ATGCGTATGTGTGCACACATATGTGCGCATATGCGTATGTGCGCACACCGTGCGCATATGC
 GTATGTGCACACATGTGCACATGTGTATGCGTGCATATGTGTGCATATGTAAAGTATGCAT
 GTACACATATGTACATATATGTACGTATGCATGTACATATGTTTGTGTATACACACAA
 AAAACCCATACATACATTCAAAAAATGAAAAGACCCATGTCTGATTCTTCTTAGTATCCCT
 AAACACTGATTCTACTCAAGAAATTCCTTTGGAGCTAGTAGTCCAGAAATCCAGCATGTGG
 AGCTGAATATGATTACCTTGGGAATGACTAGTAAAAAGGAAGGCAAGGTGCGATGATAAAA
 GATTCAAGCTGAGATAACTGGGGAGATAAAATCTGGATCAGGAGTGCAGATTTCATGAACA
 CCATACTGTGCCCTACTGTCCATGGCTTACATGGTAATTGGGATGTATTCCTAGATGTGGT
 GTGGTGTCAACCAAGTGTGACCCAGGCATACTAGTTTACCCATCGGATATGTTGAAGTACCT

GACCTGTGCTGACTCAAAAGTTCAGAGAAATGGAACTGTCTCATTTCTCACTCCATTGT
 ATCCAGTGTGCTGCTTGGCACACCAGTAGGTAGTCAGTAAGTATTTATTAAATGAATGAC
 ATACTTAATTTTGTCTCTACTGCTAGTTTGCCCTGGCGCATTTCTTTTATTCTCTATTA
 GCCAAGGCAGGGTTAGGTGTTTGTGTTTGTGTTTTCCTCATATTTTGTGAGGATTAATG
 ATAAAGCACTTATCTTGTGGAAATGAGAGTCTACTAACAAATCCTGTACATCATAATAAC
 AGAACTAT

SEQ.ID.NO. 452 HG1015033N0 PLT00014333D07.contig
 GGAGCGGAGCCGAGCCTCGCGGCCGGGCACGGCGCCGCCCTCCCCGGCAGTG
 CCTCCGACCCCTGCCCTCCGCTGGGACATGGCTGAGGCCCGCCCGCCCGGGAC
 CCTCCCCACGCTGCCCGGACCCTGCCCCAGGACCGGGGTCTCTTCCCGGCCCC
 GCCCGCCCGCGCGAGAACAGCCCCGGGGCAAGCGGGCGGACCGAGGGAGGCC
 TGGCCCGCAGCGGGAGAAAGGTGCGGGCAGCCCCAGTTTCCACCTTTTCTCCTGGC
 CCAGACGCGGCTGGGGCGGACGGGATCTCTCGGCTCTGCCCTCTTGCCTCATGG
 AGCCATGCGCCTGGGTGAGGCTCCCGAGAGAAAGCTGGCCTGCGGGCGGCCGACGAGC
 TGGCGGACGGGGCGGGCGAAGGAGGCCCGCAGGCGGAGGAGAAAGCGGGCGCGGAG
 GCGGCGGCGCGGGAAGAACTAGAGGTATTTCCCGGGCGGCTGGAGGACTGATCGAGC
 CGGACCCGAGTCTCCGGTATCCCAGCAGCCACCGGAGGCAGTGAGATAATGGAGGAAG
 AATGTAGGGAGACCTCCAGTGACCGCACTCCTTCTAGCGAGCCGGGAAACCATAGAGA
 TCAGGGCTCAGCCGAGGGCCGCCCACTGCTGTCACTGGCCTCCATCCGTGGCTTTA
 TTGGCTGGTGGCTTACGGGGTTGAATTTTGGACGCTGCCTGCCATTCTGTGGACCG
 AGTTTCAGGCAGCTGGAATAAGAGAGACATCGGGGAAAGCCGCGAGAGGCCAGCGTTC

CCTGGCTGGAGCAGAGCTAGCCGAGTAGGGCGCCCGGCTGTCAAAGTGGCCGGCGCAGT
 GCACGCTGGGCGCCCGGAGCGGTGCGCAGAGCCCGTCGGGAGTCGTAGTCCGGGACGGG
 CCGCGGCATTGTCCGCGACGCAGCTCGGGATCCGGGTGCGGGTGTCCGGCCGGGTTGCTG
 CCGGGCACCGTCGAGCGTTAGCCACCCAGCATTGAGCTGCCAGCGGCTGTTCTCCCTAAG
 CACCCTCGCTCAGTCGCCCTCGCCTCGCCTGACCGGCCGACGCCTTGGATACCCAG
 CCTTAATCGAGCCGACTACGGCCCCAGCCCCCGCGCGAGGTGCGCGGGGTTGCGGTG
 CCAGCCGCTCGCCCCCTGGCGGGGCCAGCCTCTTGATGCACCGGGGACGGGCGCCCCCATC
 CGACTCTTCGAGGGCGGCAGGGCCCCATCTGTGCTTTGCTCTCGAGCCCCCAGCTAG
 AGTTGGCTTCAGCGGAATACCTACTGTGCGGGGATTATTCAACAAGCCGATTGATCACATT
 CTTCAGCTCTAGCAGTGCAAAGGCTTCACTGAAAAAGAACATTACAAATTTTTTTTGAAA
 GAAATTTGGGGGTAAATTTAAATTTAAATACTGATTTTTTGATAACATTTGATCTTATGAC
 AGCATTTGTATTTTAAATTACAGATAGGATTACACTAAAAGTCATTTTCTGACCTGCTCCC
 TTACCCCTTATGGCTAAATAAACTGCTAGGCCCTGA

SEQ.ID.NO. 453 HG1015035N0 PLT00014333D15.contig
 GCCCCAGGCGCACTGCGAAGCGTCTCTTACGAGTCTGGCGGCCGACTTGGCCCCACGTG
 GCGCGGTGTGGCGCCAGAAATTCAAACAGCTGCAATTGGAACGACGCTTTCGCCTGCCA
 CGAGCCAAGTCCGTCCAGGACAGATTGGGTCCGAGAGTGGACACACAAGTTTTGCAGCGCAG
 GGCGCCTCAGCTTAGGACCGATTGAGAAAGCTGGGCTGGCTGGCTGGCTGGCGCTGGGC
 TGTAAGTGGCCCCCAGCGGAGGCCGCCGAGGAGCGGGCGAGCCCTGGCCCGCAGCACTCC
 GGGAAAGGGCAGTTGGCTGGCAGAAAACGCTTGGGTGAAAAAGCGGAAAGCGCCACGTGAG
 AACAGCCCCGGCGTGGCGCCCTGCAGGGCCTCAGGGCGGGCATCAAGGCCCCCATGGGGAT

CCATTCCTCCTCACGCTTCCTCGTGCCAGGCATCCGGTTACACTAAACGTGACCATACAA
GTCTCCCTCAAAACAGCGAACGTGAGGTTCAATACTCCATTTCGCAGAGGAGAAAACTGAG
GCTTAAACAGCCAGTCTACAGGATATCAGAAATTCTAAGAAAAAACTTGGAATCATGTTC
AGGACTATACAGGACTTTAAAAGCTTTATCTTCCCAAACCTGAAAACCCAGACCTCTCAAGT
CTAGCTGCTCTCCTCCAGGTGCTGCATAAATGGACCTGGTCTTTGGCTACCTTACTTTGG
ATCACTTGTTGAAACAGTTCCTTAGTCAACTGTCCACTGGACTTAAGCTGTGGGACA
CAGAAATTTTCCCTGGAAGAACAGAAACAATGAAGATGAATAGTTGAAAAATGAGAGAAATC
CTGCAATTCTAAAGGATCTAGTCTGGATGCCAGGAGCTGTGTATAGGATGAGAGCCCTGGTG
GACCTGATTGACCGCCTGTGTTCAGTTCAGTTCAGAGACTGAAAATCAGGATTTGTGCT
GGAGAGTCTTTGAACTGTGTGTGAAAGTATCTGAAGGTCCATGACAAATTTGAAAAGGTAAA
CAGTAATGGACTTTTCATAGATCTATAGTATTTGCTCTTGACCTTGGCTGAAAATTGTAAAA
CAGGCTGCCACATCTCTGTAAAGTACTTTTCTCTTCTTGCATACAAATGAATTGGTTTTTT
ATAAGTGTCCATTCAATACTGTAGCCCTTTTGGCATCTGTCTAGTGACTTTTATGGGAA
AACTTTTGGGGGATAATTGATTATAGGAAATTTGTCAATTTGTGCACAAATGTCTGAAGGGT
TTAATGTCTAATTTAACAGAAAGGCATATGAAAAAGGTGCAGGAATCGAAGGCACGAGATA
CACCTGAAAGTGGTGAGCAATGAATCTGAACAGGATGATGGATTGAAAAGTGCTCAAAAG
AGGCAGCTAGGATCCTGTATTCAATGTCTACCCTGAGCCAGGAAAGATCCTCCACACA
TCTATCAGAAAGACTGGAGGATTATGTCTGGAGAGATTGAATTTAGGGATGCTCAGGCCC
AAATGAAGGTGGATGACAAGCCTTGATAAAATCTTTTCAGTAGAAAATAGTAGAGAAATTTT
AAGAGAAATTTTATTGAAAATTTTGTAGACATTTTGTATGTCTGGGAAAAGAGGAAACTAA
CTTGTTAGTGCTTTTCTGATTTGAATAATATGTGTGCGGAAAGAACTGGAGGTGAAAATGTG
AATGATCCTTTTCTGGCTTGTCTGGAGGTTTTTGCCTCAACAGAGAAATCTAAAGGGTTAT

GAATTTATAACATTATTTCTTTTGCTCTCCAAAAAATGGAAGCTTTTCATAAAAAATACATCA
AAGTTTTAGAAAACAGCATAATTTGTCTCTGTCTTTCTTTTIGGAATGATTGTTATTGG
AAGTTAGACCCCTCCTTGGTCAATCATCTCATTTTATTATTATTTTCTTATTCTTGT
TCTAGTCTCTGGCAGATTTTCTCAACTTTATTTTCAACCCCTCTTTTGAATCTTCACT
TCCTATAATTGTATTTTCAATTTCCAAAGAGCCCTTTTGTGTTGTTCTCTGATGGTTC
TTTGTTTTTTCAGCCCTCTTATTTTATTTTCAATGATACAATACTCTCTCTTTTACAA
GGAGATTAAATTATTAATTTTGAAACTTTCTTGGCTATTTTCTCTGTTTCTTGGATTGTC
TTTCTTTTAAACATTTTGTCTCTCTTTTCAATTTTGAATCCGTCTTTCAAGAGAGGGCT
ATTCTTTAAGATAGAAAGTAATAACAAACAGACTAGAAACTGTATTTGGATGGAGAATGG
ATATAAACTGGTAGGCTTTACTCGAGAAAGGTTCTGGCTGGGCCCTTTTGTGGAGAGCCA
TCAACTTGCAGTATCTTTAGGTATTTTGGAGGGTTGGTAACTTTTCAGAGAGACTC
CACTAATCTTTTGTCAAAAGAACTCTAAGCCCTGCTTGCAGGTTCTGCCAGCTGACTGGGGA
CTAGGGCTTGAGAAATTTCTTTTAGTTAAATACCCCTGTGTTTTATCAGGGAATCTCGTATTC
ACCTTAATCTGTGCCTGCATCCGGCATCCGCTGAATTCAGTTTTCACATAGTAATCCTC
GAGACTCCTGGTATGATATGGGGTCTGTCTGCTGCTGTGATGGGTGAAAGGAGCCC
TGGAATTCTATCTGCCCTCTTTTACACACTTTCAATCCTATTTTATAGATCCTACTCTGCT
GCTTTCAGAAAGTATTTGGTACCTTCATCCCTGAGGTGTTATACCAGGGTTTGGCAAAC TG
TAATCTGCCAACCAAAATCATGATTTGCTTTGTCTTTTCAATCTTTGCTGTTTATATAA

SEQ.ID.NO. 454 HG1015037N0 PLT00014333E01.contig
GCACTGATGGCAGCTCCTTCTGTGACGGGCTCAGGAACACCTGTTGCTGCTCTCCATGA

GCTGTTCCAACTGTGAAGCAAAAGCTTCATGCATAAAGGCATCCCCACGCAGTCAATGAC
GCACAGAGAAAGAGACAGAGATAAGCTTCTTGCCCTGTTCAGACACTGAGGATCAGGC
AGAGAGAGAGGCTCCTGCCAGGCTGAGCTGTGCTCTGAGGAACATTTCCGGAGCCTCCAG
AACTGAAAGAGGGGCCAACTCTCTGGAGAAAAATGAATGGATGTAGGCCCAAAGGAGCCG
AATGTTTTCCCTTGGCTGAATGGCAAGCAAGCACCAAGTGTCCACAGTTTCCGGAGTGTGG
ATATGAACAGCCGCTCTGGGTCCTGAGGTGGAAGCCATGTGTGGAAGATGGAGGCATCG
GTTAGAAAGGAGTCTAGTCCCTGATGGTCACTGAGCTGCAGAACCAAGCCTGGGCTGCTTCC
TGCTGGATGTCACCTACTAGAGAGCGAAATTAAATGTGCTTCAGCTACTGTTACTTTGGG
TTTTCTGTCAATTTGTAGCTGAAATAATCCTAATCAATATGAGATATATTAAGTAAACAAA
AATGCAAAATGAAAAACGATTATGAATTGTATGATTCCAAATTTATCTTTTATGTGATT
TAAGTCTTTAACTTTTAACCCAGATGATTTTATTTTAAATGTGTGTTTATATACATCCA
TATGTACATAGAAAAAAGACAGGAATGTTGAAAGCTAAAGTATTAGTAGAGATTATTTCA
GGATGGTAGGTTATAAGGATTTTCATGTGCTTATTTATTTACCTTTTGTGTTTCACTGT
ATTTGCTAATTTTCTATTTTTTTATTTTTATTTTTTGAGACAGGATCTCACTTTGTCCAC
TCAGGCTGAGTACAGTGGCACAAATCATGGCTCACACAGCCTTGACCTCCTGGGCTCAG
GTGATCCTCCCACTCAGCCTCCTGAATAGCTAGGACTTGGTATGTGCCCCGTGATCCCCAG
ATACCTGGGAGGCTGAGGCAGGAGAAATTGCTTGAAACCAGGAGGTGAGGTTACAGTGAG
CCAAGAAATCCACCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTGTGTCTCAAAAACA
AAAAAACAAACAAACAAAAAACCTTATTCTCAGAATGAAGATCTGAAATAAGAAAC
AAACAAACAAACAAATAATTTTAAAGGAAATAGAAAAACAAATAACTTATTCATGAT
TCCATGGGAAAAAGACTGAAAAATAAGTGTTCTTAAAGTAAATAAGGGCATGAGAAATTGT
TCAAAATATAGTGTTAAAGAGAAAAAAGGTTATTCAACAATAATGATGCCATTGTTTTTGA

TAAAAAGAAATTAACAACATAAATGTATGCTAGAAAAAGAACTAGAAAGGAAATAAGCTGA
TATATTGGTGGTCTGGTCCAGAGGAGAGATTTTAGTGATTCTTATTTTCCCTTTATAC
TTTTGGTATTTTCTAAATAATTTCTCTAATAAACCTGTGCTAATCTTATTTATTTCAAAAT
GTTTCTACTTAGAGAAATTTTAACAAGACAGAAAAAGCATGTTATATAAGAAATAATAAAA

SEQ.ID.NO. 455 HG1015039N0 PLT0001433E04.contig

TAATTGCTATTCTGTAAAGACATACAGTCTGTGTAAAGATGTATCTTATTTACAGAGACATT
TTTGAAAAATTAAAAATATTAAATACTTTTGTGTTATATAGAGACAATGATCTGGAGGTATAA
AAAGAAAAATATTATCTTGTGATGTAATAATATGATATCCTTATATATATAGAAATCCAAT
AAGATATCATGGCGCAATATTAGCAAAAAAAGTCAGTCTTACCCTGAAAAAAAATAG
ATATGAAAAATCTCAGCACTACACAGAAATGTTTAGATGTCAATCAGTTTGTGATGCAGC
ATTATTATAGTAGTAACAAATTTAAATCTAAAAATTTTAGAGTAATACTTTCCTTCTGATC
AGTTGCCCTGTGGTTATTTTGTGTTAAGGAAACTACCATTTATGACGGGAAACTCTTATTTT
TGCTGTTATATTTTGTTCCTTATTTCACCACTAGAGGTGACATCATTTATTTGTAACTT
TAGTACTAGGCTTTCCATTTGTGGTATGGACACGTATGATATGGAGCAAATTTTGAATTG
TCTTTCACATAGGGGAAATTGCATCTTATATACCCATAAATTGGGTTTTAGTGAAAGAGC
CAGGATGGAACATAAATTCCTAACTTCCATCCTAGTTTCCTTTCTACTTGTTCAGAAATGG
CCAGTGAAACACACAAATGGCACTCAGATTTGTGTATGGCTCACCAACTCAAAAGAGGTTT
CTTCTCTCCTTTGATAAGTCCACAAAAACATCTACTGTAAAGACTTGAAATACAATGCC
TTTCTTTTTTATGAGAGCCAAAGTTTACCCACAACCTCTCAACTACCAGTACTGGTCTGACT
GTGCAGAAAAAGGGTTGCCCGCAAAGGTCAATGATATCACCCCTAGAACATTCTAGAAATGAT
GAATTTTGAGAAATGTGAACCTTAGAACCTCGTGGGTTGTCTGCCCTTGGAATTCCTGTG

CTGTATCCTGAAATTTAGCTGAATTTGGCAGAGAATTCAGAGAAAATTTGCTTTTCTAA
 TCCTTGCTGCCCTGAAAGTGCTAATTAATTTTCCATAGCTAGGATAATTACCATGGAGCAATA
 TTTTCTCTGCCCTTTGATAACTCCCAGCAAATTAATACTAACCGCCTTTTGAAATCCAGG
 AGCCACAAATTTTAATTCAAGTGGCCTGCTAAACATGCACCTCTCTGGTTTAAAGTTTGCC
 AGAGGTGGTGTGCCCCCATATCCACTGGCGCCGGCCTCTTTTCTTTCTTTCTTTTCTTT
 CTTTCTTACTCTGTCTCTGTCTCTCTTTTCTCTCACCTCTCTTTTCTCTCTCTCTCTTT
 CTCTCTTTCCCTCCCTCTCTCTCTCTCTCTTTATTTTCCTAAGCAAGCTAACACAGG
 ACATAATCTCTCTTTTCTTTTGTGAGATGGAGTCTCCGTCTGTCACTAGGCTGG
 AGTGCAAGTGGCACAAATCTCAGCTCACTGCAGCCTCCGCCCTCCAGGTTCAAAACGATTCTC
 ATGCCCTCAGGGCTTCCGAGTAGCTGGGACTACAGGTGTGCACCAACCATGCCAGCTAATTT
 TTGAATGTTTAGTAGAGATGGGATTTCAACCATGTTGGCCAGGCTGTCTCAAACCTCTGA
 GCTCAAGTGATCCACCTGCCCTTGGCCTCCCCAAAGTACTGGGACTACAGGCATGAGCCACT
 GTGCCCAGCTGACATAATCTCACTTAATAGAAATATAGAGAAATTTTTCATGTTTTCTT
 TTTCTCTCCCACTTTTTCATAATTCCTCTTTTTCATTTTGGCCTTTCCGTTTCTGTCTATG
 ATGTAGGCTTCTGAGGAGAACCAAGAGCTTGGCTTTAGTGGTAGAATGACAGAACTTAG
 GGATCCCTTGCAGGCTAGAACAAAGTTCTGACCCCTTAGACCAAATCTTTATGTTAAGAGT
 TTTCCAGAAATTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
 ACACACACTGCCCTTAAAGGAGGAGCCAAATATTTGTTGTACCTGTGAACCTGAGGAATTA
 TAGATAAACCTTAGGTCAAATCATTTTCACAAATTGCAATTGGTGGTATTGAAAAATGATGAG
 ATTTCTCTGACAGAGAGCTTTGTCTCTAGTTTGTGTTCTTCATAGGTCAAAACTGGCAATA
 TTCTCTTGTGTGCAAGATAAAGTGTGTGTGTCTCTATCACCATATGCATGAACATGTAAAG
 AATCAGATACAAATTTCTGCTTCATCAGTTTCACATGTTTCATGTGTGTCACCTGAAAAAATGC

ATCTACTGTTTATAGCTCCCAAGGAGACCCCAAATCCCTTTTCTTTTCTTTTGAGATGGAG
TCTTGCTCTTGTGGCCAGGCTGGAGAGCAGTAGCGGATCTCAGCTCACTGCAACCCCC
ACCTCCTGGGTTCAAAGTGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGGATTTACAGGT
GCCCCGTACCATGCCCTGGCTAAATTTTGTATTTTAGTAGAGACAGGCTTTCACCATGTT
GGCCAGGCTGATCTCAAACTCCTGACCTCAGGTGATCCACCCACCTCGGCCCTCCGAAAAGT
GCTGGGATTACAGGAGTGAGCCACCGCGCTGGCCCCGTATTCAATTTTATACTGGAAA
ACATTTTTTTTGTAAATTTTCTTTTGCAAAGAAATGAGCATAAAAATGAATACTCCAAAAGA
AAAGGAATTATTATGGCAAATTAAGGGACACGGTACTATGATTTAATTACTGTCCTGT
CTTTTTAAATGTCATCTCTTGTTTACCCCTTTTTTTT

SEQ.ID.NO. 456 HG1015040N0 PLT00014333E05.contig

GAAGATTACCTTTATGAATCATATGTCCCAAGGTCTATTGACTTTTCATATACTTAAGAAAG
AAATTTTCTGGCTAATTTTAACTAATTGAAAATCTTTCACTTTATGCATTTATTTAAAGT
GGCTTGAGCAGATAAGAGCTCTGTAAATTTTAAATGTTCCCTTCAAAATATAAGTTGATTTT
AAAGGTATGCATTATGATATTATTTTATTGTATAATGCAATGAAATCAAAATATATTAA
GAATTTCCCTATGCAGCACATTTTATATTAAATAGTACTCATATTAGAAATTAACAACAAA
AGCAAATACAGAAGCATTTACTTAGCTCTTACTGTGATCTGGGCACTCTGCTGAGTACAT
TGCAGGTATTATCTTATTCAAATTTACCCCAACCATAGTAGATACTAAAGAGAAATGCATGGC
TTGATCCAAGCCCAAGGAGTATGCCATTGAATGGAGGAATGAGTAGTGTACAGTTATG
TTATTCTTTTCTTTTCTCCTTTTGTGAGACAAAGTCTCACTCCAGCCTGGGCGATA
GAGTGAGTCTCGTCTTAAAAACAACAAAACAGACAAATAAAAAACATACGATGGAGCT
CCAAATATGTCTGGCAGCAGACTTCTTAGTGAAACCTTGTAGGCCAGAAGAAAGTAGCAT

GACATATTTCAAGTACCGAAGGAAAAAAATACCTTGTATCCTAGAAATAGTATATCCA
GCAAAAATACCCCTTCAAACATGAAGATAAATAAGAACTTCTTAGACAAGCTAAAAGGAT
TACATCAATGCCAGACCTGTCTACCAGAAATGCTAAAGGGAGTCCTTCAATCTGAAAAGT
AAAGGGCGTTAATGAGCAAGAAATAAAGCAATTGGGAGGTTCAAAACCTCACTAGTTACAGTA
AGTACACAGAAAAAGCCAGAAATATTATAACACTGTAAATTGTGGTGTAAACTACTCACG
TCTTGAGTAGAAAGACTAAAGATGGACTGATCAAAAAATAATACTACAAGTTTTCACAC
ATAGACAATACAAATAAGCTATAAATATAACAACAGAAAGTTAAAAAGTGGATTACAGAA
AGTTAAAAAATGAAGTTAATGTATAGAGTTTTTTTTTGTCTTCTTCTTATTGTTT
ATGCAATCAGTGTTAAGTTGTCAATCAGTTTAAAAATAATGGGTTATAAGATATTATTGCA
AGCCTCATGGTAGCCTCAAATAAAAAATCATATAACAACAAAAATAAAAAATAAAGA
AATTAAAAACACCACTGGAGAAAAATCATCTTCTACTAATAAAATGGCAGGAAGGCAGAAA
AGAAGGAAGAGAAAGACCAAAAAAGCAATCAGAAAAACTAATAACAAAAATAGCAGGAGTAAGC
CCTTACTTATCAATAATAACATTGAATGTAAGTGAACTAAACTCTCCAATCAAAAAGACAT
AGAATGGCTGAATGGATAAAACACAAAGACCCCAATGATTGGTTGCCTATAGGAAAATACACT
TCACCTATAATGACACACATAGGCTGAAAATAAAGGGATGGAAGAAGATGTTCCATACAAA
TGAAAAACCAAAAAGAGCAGGAGTAAACTATACCTATATCAGATAAAATAGATTTCAAAA
TAAAAACTATAAAAAGAGACAAAAGGGGTCAATTTAGAATGATAAAGGGGTCAATTCAGC
AAGAGGTTATAACAATTGTAAATATATATGCAACCCCAACACTAGAGCACCCAGATATTATA
ATAAGCAGCTAATAACAAGCAAAATTAATAATAAAGATTGGAGCAGAAATAAATGAAATGG
AAATGAAAAACAATACAAAAGATCAATGAAACAAAAAGTGGGTTTTTTTGAAAAGATAAG
TAAATTGACAAAACCTTAGTCAGACTAAGAAAAAAAAGACAAAGACCCCAATAAAATTAGAC
CTGAAAAATGGAAACATTACAACCAATGCTGTCAGAAATTCAAAAGGATTATTAGTGGCCACT

ATGGCTAACTATATGTCAATAAGTTGGAAAAACCTAAGAGAAAATGGATAATTCTCTGGATAC
 CTACAACCTACCAATATTGAATCATGAAGAAATTCAAACTTGAAACAGACCAATAACAAG
 TAATGAGATCAAAGCTGTAATAAATATATAGCAAAAGAAAAGCCAGGGCCTGATAGCTTC
 ACTGCTGTATTTTACCAAACATTTAAAGAACTAATACCAATCCTACCCAAACTATTCTGA
 AAAACAAAGGAGAAAGCAATAGTTCCAAATCATTCATGAGGCCAGTATTACCTGACACC
 AAAACAAAGACACATCAAAAAAGAAACTACAGACCAATATTGTTGATGGATGTTGAT
 GCAAAAATCGTCAACAAAATACTAGTAAATTGAAATTCAACAAAACATTAAAAAGATCATT
 TATCGTGATGAAGTGGATTTATCCAGGGATGCAAGGATGATTCCATATATGTAAATCA
 ATCAATGTGATACATCATATCAACAGAAATTAAGGACAAAAACCATATATGATCATTTTCAATT
 GATGCTGAAAAAGCATTTTGATAAAATTCGACATCCCTTTGTGACAAAAAGCCCTCAAAAACT
 GGGTATAAAAGAGACATACCTCAACACCATAAAAGCCATATATGACAGGTCCACATGAAA
 TCCTTTCCTCTAAGATCTGAAACACAACAAAGGACACCCACTTCATAAAGTTACTCAACA
 TAGTACTGGACATCCCAGCTAGAGCAATCAGACAAAGAGAAAGAAATAAAAGGCACCTAAA
 CTGGAAAGGAAGACATCAAAATTATTCTTGTGTAGATGACATGATCTTATCTTTAGA

SEQ.ID.NO. 457 HG1015042N0 PLT00014333E14.contig
 GCCACTGGTAACATCCAGCCAGTCAGCAGTGAGCTTCTGGGCAATTCTTTCAAGAGACTA
 GATTGACACCTGGAGCCAGACTTTATCGGACTTACCAAAGTTTGACCCCAATAAACAGTT
 CAGTTTCCAGAGTTTCTGGCTGCATGTAGTTGAAAGCAGGACTCTGCCAAGTTCTGAG
 TCCAGCTCTGAGCTGTTGTCTTCCACCCTGCTGCAGAACACGAGGTGAGGAAGTGGA
 GGCTCCCAAAATGTTGACTTGTCAACAGGTGACATGGTTTGTAAAGTGGCAGAACTGGGGCGA
 GAAGCCAAAGTGGTTGCCTCCTGTGCCCTTCCTTCCATTGTACCACAGCGCTGGCCTCTCCT

CTCCTCACCTGACACTCCATGGGTTCAAGTCAATTGCTACCGGATGCCAAACTCTGCAGGT
GCTGTGCTGGACACCGGGAAGCCAAGTGACTGGATGAGAGACACTGTCCTAAGAGTACTT
GCCATCTAGTGCA TGACACAGACATCCAGTGAATAATAGCAGCACAGTATCATGAAGGCT
ATGGTAGGGACAA TCCAGGATCCTACCAAGAGCTTATTCACATGGGTGGTCAGAAAAAAGAAC
GGCTGCCCCATGGCATTGGGGCAGAGAAAAAATTGTACAACAAGCCAAAGAAATGCCTATCC
TTTTCATGAAGCA TTTTGTGGGGCCATTCTGTGACCTGTCTATGCTTAATGATTCTAGGA
ATGGATTCA GTTACAGGCTGGACTCAGGGCTTCTTTGTCAGTAA GTGACAGGCTACCCACAT
ACATACCACATA TGACACACTGTGTCTGAACAAGAACACTCGCTTAAAGGAAAGTCAGG
AGGAAAAGGAAAGGTGACTTTTACTGCTTTGGCTGAAGGTACTGTATTTGCCAGTAGAG
AACTTGGGGA AATCAAATATTCAATTACCCTATAATCATGGAAAGGTAA GACTGATGGGT
GAGCTTTTAGCTGTCTTCACGGTGTTTACACAICTTTACATCCCATGGACATCCCTTTGG
ATCTAAATGGA AGTCCTAATTCCGAGGTCTTGATTTATTTTAGACA TGGGAAGTACTTCT
TAACTGAACAGTCAGGGTTAAGTCAGGGAAGGGAACCCCTTGGAGTTATGGGATAAAGGA
CACACTTATTATAGGAATTGGACATTACACA AACTTTGGGAGGAACTGGAGAGGTGAAGGC
CCCGAGGAATGGAGGATCAGAGAAAGGAGCCGCCAAGCAGTCCCTTGGACAAGCTAAGC
ACATCCTGCTGTGGAGGGGCTGAAAAGAGAAAGCCTGTGTGTCCTCCTCCGTGGAAAT
GCCATCTGAGGAGGGGCTGGGCTGCTGTGCTGCTGGAGTCAGCAGTCAGGAAAGAGGA
CTGGATGTGGAGTGACAGCAAGGACAAGCTGGAACCCGCTGGCCTCTCTGTGTCTG
TCACCACATCTAACCCCTGATGACTTTTCTGAGAGTGATGGCCACTGCTTCACATTTTCCAA
ACCTGGTACAA GTTCTTCTTGGCCAAATTTAACTCAGAACTGTACAGGGAAAGGGATC
TCGAAAAAGATAGTTCCATCTTAGCCAAAGTGAACAGAGCACA AACCATCACATTTTGAAAGAA
AAGGCTGAAGGTAGAGAGAGAA TACTCTGCCTTAATTAAGAAACA AATACTTGGCTGGGT

GCGACGGCTATGCCTGTAATCCCAGCTCTTTGGGAGGCTGAGGCCGCGGATTGTTTGG
 GCTCATAAGTTTGAAACCCAGCCTGGCAACATGGCAAAACCCCGTCTCTACAAAATA
 CAAAAATTAGCTGGGTGTGGTGCATGTGCCCTGTAGCATCACCTACTCTGGAGGCCGAAG
 TAGGAGGATGGCTGGAGCCTGGGAGACAGAGAAGTTGCAGAGAGTTGACATCACGCCACTGT
 ACCCCAGCCTGGGTAACA

SEQ.ID.NO. 458 HG1015086N0 PLT00014333E15.contig

TGAATGTGTATGTTGGCTGCAATTTAGATAATGTATTACCTTCATGACAGTTTCCTTTC
 CCTAGGGAAAATTGCCAGGAATCAATTAACCTGAGTGTCTTAATGGAAAACCTGGATTGG
 AGGATACTTATAATCCCCCCCCACCCCATTCAGTGTCCCTTATATTGGGAGGAGTCAGTG
 GAAACCTTTGTAGGGGATCTTTGTGCTTTGGAGTTTATCTGTGCTGTTACTATCTGGA
 GTCAGTGATTGTGGAAAGAAACATGTGATCCCGGTGTGCAGAAAGCATCAGAAAATAGGCT
 GTTTCATTTACTTCCCTTTGGGCTATATAGAGAAAATGAGGAAAAGTAGGGATTTTGA
 TTTTCTGCAGATAAAATGAGTTGAGGTTTTTAGTATGACTCAATAGATCTTAAAGTGTTT
 ATCTGCAGTTTCTGTACTGCCTAAGATACATTATAAAATCTGTATGTAAAGTTTGTAA
 GAGGAAAGATTATGCTTATCTTTTAGGGGAAGAGGGAATATTGAGCTACATTTATAGAGT
 GATTTTGTTCAGGTTTTTCTGAGGAATTTTAAATACCCGTGTCTCTTAATAGTAAAA
 ACAGTTTGAGGTAAGTTTCTTACAACTATTTGTCTCTGAATAGTTTCTCTTAAA
 ATTTGACAAATATTAACAAATGTCTCTTCTTCTTTGGTTACAGTGGAGTCTCTCTGT
 TTTGCCAGGAAAGTTAACATATTGTTTTTTTTCATTTTTTGTTCATCATTTACAAATTAGAT
 GGTATAGATTTTAGCAAAATAGATCTCAAGAACTGGATTCAATTAGCAGCTTTTAAGGACA
 TAAACATTAAAGGGTTAGACGGAAAAGAAATCGTTTTCGACCGACCGGATAAACGACAGGGC

CGTGTTCACGTGTGTCCAGGAATCGGTCTGGAGCTGGTGCCCAACTTCTCAGTCCATCT
TGAGAGCCAAATAAAATACTTAACCTAGGCCTCATTTTGTTCAAAAAATAAAAAA
GTTGATTCTTTGTGAAGACCTTAAGAAATGTTTGGCATACCTTTAGACTAAGTGACAGTT
CAATTAAGATCTGGGTCCCTAATAGAGGAAAAAAATCACTTAAAAAAGACAGATCCCA
CAGCGAATTATATTTGGTTACAATACACTTGGGCTATAAGATGTTTAAATTAATTGT
TCTTTGCTTTAAGTTAATAACATATTAGTATTTTAAATTTATAGAGTACCTTTGAATTC
AGATGCTGTATAACTTCAGTTACTTGGTAATAACCTTATTGTGGAATTGCCCTTTTACT
GAAAGGTATCAACAATTTCTTGGTGGTGACACACACACACACACTCTCTCTCACAC
ACTCATACACACTCTCTTTCTGAAGAAAGGAAACAATTTAAACCATGAGATCTCTGCCCCA
CAATTGATATTAGACGTTTCCACATACTATTCTTAAACCACATATCAGTGAAGTGTGG
TGTTTGTCAACGCTAACAAAGGTGGGCTAATTTAATAATTAATCCCAAATTTTAGATTC
CTCAGGCAATTGGGGGTTAGAAATCCGATTTTCATCCATCAAGCTACTTCTGTTCTTTATGA
CACATTAACATATGATCATTAATAATGGGCTGTGGTAAACACTATTAACTTTCACATTA
TTTCTATAACAAATTGATGACAAATGTATAATTTCTATAATCATTAATGTTTATATGACTG
TCTTATACAGGCTTTTCTTCCAAATGGTTGGATAGAAAAATATCTTTTTTTTTTTTT
TGAGTTGGAGTCTTCCCTCTGCTCAGGCTGGAGTGCAGTGTGCAATCTCTCGGCTCA
CTGCAAGTTCGGCTCCCAAGGTTACGCCATTCTCCTGCCCTCAGCCTCCCGAGTAGCTGG
TACTACAGGCCACCGCCACCAAGCCCGGCTAACTTTGTTTTTGTATTTTAAAGTAGAGACG
GGGTTTCAACCGTGTAGCCAGGATGGTCTCAATCTCCTGACCTCATGATCCTCCCACCTT
GGCCTCCCAAAAGTGTGGGATTACAGGCATGAGCCACGGTGCCCGCCTAGAAAAATATT
CTTTACAGTAAAAAATTTGGGATAATTTTTAGACCATTATTGTTTCTGAAAAACCAAGG
CTGGAAAAAGCAGCTTAGTGTAATTTCTAATATTTTAAATTTCTTAGGTATGGAGAGTAG

AATAAGATTACTCCTCCAGAACATTTATTTTCAGACATTTTATACATTTAATGTTTTTGT
 ATGCTGTTTTCTCATGTTATTAATTTCAAAGCCCTTGGCAATAAGGAGGTTAAACTTGTCT
 ATAGTGTAAGGAAATTTAACTTTAAAAATTTTCCTTTAGATATGCTAGCATCTTAATCTT
 GGTACACTACATATTTTATCGAGTCAAGGCTTTAAAAATTATAGCTTCATTTCTGTTTGAC
 AGGTTTTTAAGTATGCCTATACACCATAAATGATTTTATTTTATTTTAAAGATGGAAATTT
 GTTCTTGGCTGAAGTGTAATGGCCCAATCTTGGCCCACTGCAACCTCCGCTCCCAGGAG
 AATCACTGCAACCTTGGGAGGTGGGGTTGCAGTGAGCGGAGATTGCACCACCTGCACTGC
 AGCCTGGCAACAAGAGTGAAACTCTGTCTCAAAAAAAGAAAAAGAAAGATATTA
 GATATTATACTAATAATCTTAAC TTGCTTAGAAGTATGTGAGTACAGCTGGGTGTGTGG
 CTCCTCCTGTAATCCCAGCACATTGGGAGGCCGGGGTGGATCACAAGGTCAAGAGAT
 CAAACCATCCTAACCAACATGGTGAAACCCCGCCTGTACTAAAAATATAAAATTAGCT
 GGGTGTGTGTGCACACCTGTAGTCCCAGCCACTTGGGAGGCTGAGGCAGGAGAAATTGC
 TTGAACCCAGGAGGAGGTGATGGTGAGCCCAAGATGGTGCCATCGCACTCCAGCCTGA
 GCAATGAG

SEQ.ID.NO. 459 HG1015044N0 PLT00014333E24.contig
 CCACGGAGTCAGTGCAGGGGCTATGAGAAATTGCAATTTGCATCAAAACGGTAAATCAAAG
 TGAAAGAAGTCATTAAGTCAGGTACAAATACTTATCAAAGTGCCTGTTGTTGCTGGACA
 TAGCAGGAGTGCTCCCAGGCAAGGAATAGACACCTATTTTCAGACTTCTGATCTCTTTG
 GGGATTGCCAAAAATGACCTCAAGTAGTGAAGATATAGTGGAAATAATCTGTATTTCTGCTG
 ACCTTGATGACATTTGGATCTCATTTGAGGTACTAAGAAAGATAAGCAATTGGTTTGAAA
 AAAGCCCTTTTAAGAGCAAAAATGAATTTCTGCTAAAGTTGAAGCAGGAACAAATATCAAAT

TTATGGTGAAGCTCAGGTGGAAGATGATGAAATCACTGATGCTTTACAAGAAAGTTTATG
GGGACAAATGCCCTTAAAGAAATCAGCAGTTTACAAATGGATAAAGTTTGGAAAGGAT
GAGATGCTGAAGATAAATCCTGCAGTGGCAAAATTCATCCACATCAATTTGTAGGAAAAA
ATTTATCTTCTTCATGCCCTAACTAACCAACATTTAATAGCTTGGACAATAGCCAACACC
ACAGACATCTCAATTAGTTCAGCTTACACAATCCTGACTAAAAAATCAAAGTTGAGCACC
TTCCACTCAATTGGTGCCAAAACCATGGCACCCAGATCAAACTGCAGACAAAGAGTGGAGCT
TTCCATGAATATTTAAACAACCTGGGATCAAGATCCTGTAGCATTTCTCAAAGAATAGTA
ACAAGAGATGAACATAGCTTTACCAGTATGATCCTGAAGACAAAAGCACAAATCAAAGCAA
TGGCTACCAAGAGGTGGAGTGCAGTCAAAGCAAAGCAGACCAAGTCAAAGCAAAGGG
TCATGGCAACAGTGTTTTGGGATGCCCCAAGGCACCTTGTGCTGACTTGCTGGAAGGCC
AAAAAACGATCACATCTGCTTATTATGAGAGTGTCTGTGAGAAAAGTTAGCCAGAGCTTTAG
TAGAAAAATGCCTGGGAAAGCTCCACCAAAAGAGTCCTTCTCTACCATGACAATGCTCCTG
CTCATTCATCTCAACAACAAGAGGAAATTTGTGACAGTTTCAATGGGAAATCAATAGGC
ATCCACCTTATAGTCCTGATTTGGCTCCTTTTAAACCTGTTTTGTTTCTTAGTCTTTTTT
TAAAAAAATCTTTAAAGGGCAATCATTTTCTTCGGTAAATAATGTAGAGTGCATTGACA
TGGTTAAAAAGACTGCATTGACATGGTTAAATTACCAAGAACTCTCAGTACATTAGGTATG
GATTAAATGGCTGATATATCACTTTCTCATCAATTTCAAGACATTTTGTGTTCTCATAT
GATGCATCATGAAGCAAAGAATGACTGAAGTGCCTCTTTGTACCTGATAGTTTCCCAATGT
GTTTTTATAATTTTAGGTTCTAAGCTTATTTTAAATATGATGTGTTTCTCCTGTAGAAA
ATCTTGGTTGATAATTTGTCCCTTTTGCTACTTCTAGGCTTCCCAGGATATCCCAATTCA
GAGCACTTTTTATGTTAATTTCTCACTTAGAGGTGGCAAACCTCACATATATTACCTAT
TTCAATTATAAGCCCCCAAATTTATGTGAGGGGAGGCATGTTACAAACTCATAGGGAAG

ACATTTTCCCTCTTACTTAGAGCCCAAAACCAAGACATTCAATTTATGCAATAAAAAGCTTAC
 TAAGTTACTGTGCTAAAGACTAGGGATGTCATAAGAAACAGGCAAAATCTCTGCTCTCAT
 GAAAAGTTATATTTTGGAAAGAAATATCCTTGTGCTGATGGGGAATTTTCTAGTCCTCCAGC
 AGGAATCTCAGTTTTCTACTCCCTGACTCCATTAGGCTCAITGGCCTCAATTCATATATTCITT
 TATGGTTATTATAACCCATTTCCTATGATTACCAAGGCTGGCAATATTCITTTCCCTCCACCA
 TCTTCAAATTGACCTTACCATTTTATTTTGCCTTCTTCATTTTGGCAAGTAGGAGCTTG
 GCTTACTTTATTGTAAAGCTCAGCTATACATTTAAAAGTAAATATACAATTTCTTTTCATC
 ATTTCTATGGGTTTGGTGGCACAGGGAGTTACAATATCTGGTCTACCCACATTCGAGGAAC
 TAGATTTCCCCCAGATGACTAACTCTTCAGTCAAAAGTCTTCTTTTTTCCCGCCTCAAAC
 ACTTTAGATTTACTTATGCCACTGTGGAAAATTATTTCCACTGTCCTTGAAAGTCCCCCTTT
 CCCAGGTTTTCCCTCTGCTGTTTTTTTCTAATCTTTACGGTGTGCTCAAAATGTTACCTC
 TTCAAAGAGGCITTCCTGGAAAACCTCACCTTGAGCACTCTATCTCACTCACTAAATTTCT
 TTGTTTTTTTTTCTTATTTATAACTTTATCAGTATTTTGAATTAATTATACTTATTTATTTA
 CTTACTTGTGATCATGCCCTCTATTCCACTGTAAATGCAACTTCCATGACAGCATGATCTT
 TTTTGTCTTGTATTATCATTATTTCTCTAGGGTCAGTCTACCACTACTTTTATTGTCTTGG
 CAGAAATTAGGTGTCTCCTGAGGTGATGCAAAAAGAAAAGACATTATCAGTGATGGCTTCT
 AACCAAGGATTCAAGGCTTCTGATTTAAATTCTATATTATAGCCATCACAGGGGTGAAGG
 CACAAAGTAAATGACACCAGAAAAGAGGATTTAAACACAAAATCCGAAACTTGGGATATTTC
 ACAGGATAAATTGCTGAGCCTCTTCAGCAAAATAAGTGTCTAGCAAAATCCTGTATTAGTAT
 CTTTAGTGATCATACCTGAAATTTGCTCTCCAAGGCACCATACTATACATTCCTTAGTAT
 CCTATCTGTTTTAAGCAGATAATAAAGTCAGCTAAACCTTCCCTTACTGTTCACACACTGA
 GTAAAGTGCAATTTGGATGTTTTGAAGATATTTAAATCTTTTCAGCCCAGATTTTAAGAAATACC

TGTTTGAGAGAAATTACAAGAGCTTTTAATGCTTGCTACAGAAAAGGTAAATGTCTTATCTT
TTCTGTTTAGTATTTTGCAATTCCAAATAAGAGAAAGTTATTCATAAAGTTGATAAATTG

A

SEQ.ID.NO. 460 HG1015045N0 PLT00014333F07.contig

GTTTTTCTTTTTTGATTATAGCCATCCTAGTAGGTGTAAGTGGTATCTCATTTGTGATTTT
ATTTTGCATTTCCCTAAATGATTAATGATAATTGAGCATCTTTTTCATGTGCTTCTTGGCCA
CTAGTATATCTTCTTTGAAGAAATGTCTATTCAAGTCTTTGGCAGTTTCTAAATGAGTTA
TTGTCTTTTTGTGTTAGGTTGTAAGAGTTATAATCTGGATAACAGATCCTTATTAGATA
TGTAATTTGCACATAATTTCTCCCATTTTGTGGGTTGTCTTTTGTCTTTTATTTTTTA
TTGCTTGTGTTTGGTATCATTAATTGATAAATTCATTGTTAAATTAAAGGTCATGAAGAGT
TACTCTTACTTTTCTTCTATAAATTTTAGTTTTAGCTCTTGTATTCAATGTCTGTAATCA
ATTGAAGTTAAATTTTAAATATAGTATGTGTTAGGAGTCCAGCTTCATTTGTTTGCATGTA
TCTATGCAATTGACCCCAAAGTGATTTGTTGAAAAAATATTATTTTCTGATGGTATGGTCCT
GACACCCTTGTAGAAATTCAATTGAAGATATATGTATGAGTTTACTTCTGCAGTCTCAAT
TCTATTCTATTGGTCTAAATGTTTATCTTTATGCCAGTACCAGCCTGTTTGGTTATTTT
AGTTTGTATAAAATAGACAAACAATTTGGCCGATTAAATCAAGACAAAAAGAAAGTATA
TAGAAATAATATTATGAAAGAAAGGGGGAGCATAATTGAAAAACACCATAAAGATTAAAA
AGATTAGGGACAGTGTTAACAATTTATGCCAATAGGGTTTGAAAGTTTAGAGAAAAATAG
ACAAATTCCTAGCAAAATATAAATTATCAAAACCTACTCAAGAAAAAATAGAAATCTTAT
ATAGTCACAAGGTATTAAATATTCAATTCTTCTCTAGAGACCCCAAATATATGCTTTCTACC
CGTAACTCACTGTAAATAAAAAAGAAAGGGATAATATATTGTGCAAACTCTAGTCAAAAAA

AAGTTGGAATGACAATATTAATATAAACAAGTGGACTTCAAAAGAAGCAATATTATCA
GAGAGAAAGAGGGCACTGCTTGGTGAAATTGCCCCCTGATAAAGACATAACAATACCGGAA
GTAAAAATTGACATAACTGCAAGGGGAAATTGACAAAATCCACAATTATATTTGGGATTTA
AATGCTTTACTCTTAATAATTAAATAAAGTTAGCTAAAAAATTAATTTTAGCTAAATAAA
TTAGCTAAAAAGTTAACAAAGGATAGATCTGAACAACACTAGGATAAACTTGTCTTAATGG
ACATTTATAGACCACTATAACCAAAAGTGGTAGAATTCACAATTTTGTGCAAGTACACAATG
GAATATTCAACCAAAATAAACCATATCCTGATTTATATAAAATAAACTAACAAACTGGCTTCT
CTCACTTAGCAATATGAATTTAAATTTCCCCCATACTCTTTCATGGCTTTATACCTCATTT
TCCTTTTATTGCTGAATAATGTTTCAATTTTATGGATGGGCCATAAAATGTTTATCTATGC
ACCTATAGAGGGACATCTTGGTTGCTTCCAAAGTTTGGCCAATATTAATAA

SEQ.ID.NO. 461 HG1015046N0 PLT00014333G01.contig
GATCTCTGTGTTTTTATTTTACTTGCCCTCTCAACACTGTGTTTTATCAAAAGCCAACCACTCT
CACTTAGAAACACTTTTCTCTTGTGTTTCTGACACCACACACTCTGGGTCTTCTGCTA
CCTCACTCATTAATCCTTTCTTGGGCTTCTTAATTGGCTTCTCAATCTATAATTATCAAAG
TGTCCAGGGTTCTGTCTGAGTCACCTTCTCTATCCACCCCTCCCTATGTGAGTTTACCTA
GTCCTGTGCTTTAGATACCATTTATAAATAGATGACTCCAAAATATGTCTCTCTGACAAT
GACCTCTCTGTGTAATTCCTGGCTCCTGTAAATAACTGCCCTACTTGATGTCTCTCCCATATGA
ATGTCAGACTGGCATCTCAAAACTTAATGTGTCCCCCCTCAAAATCTTGATTTTCACCCCCAA
ACCTGCACCTCTCTCAGTCTTGTGCATTTTAGTAAATGTCACTGTGACTCATTTCTGTGTTGT
TCAAGCTAAAATTCTACGAGCTGTTTTTGTGATTTATCTCATCCCTTTACCTTTAAATCCTT
CAGCAAAAGCCTGTTGATGCTATCAAAATGTATCCTGAGTCTTTGACTTCTTTCCACCTCC

ACTAAACCATCTAGTCCAGTAGTAGATGATGCTGCTGGACTACTTGAGAGAGAGCTTCCT
AACTGACCGTCCTGCTTCCATTCTTGCCCTACTGTAGTCCATAAGTGAGAGGCCACCTTT
GATTAAAGCCATTTTATAGGTGGGTGTTCTGTAACTTGTAGCTGAACCCATTCCCTAACAGG
CAAAATAATTGTGTAACGAAGCAGGATGCTAAAAACAAGACACTAAAAATATGGAATTGGC
TAACTGGAGGTTAAGGTAGCCAGTGTTAAAAAGACTGTATGTCAAACTGGAAAGTTGGTGA
CCCTTATAAGCCGTAGTGAAACTGTCACTTGCTGGACCTTGGATATAGGCCACATGCCAA
CCAATGCTATATTTTGGAGAAAATCTCAGAGCAATATGGAACATTAGTGAGTGTTCCT
TTTTTAGTAGCTTTCAGCAGACTCCTTCAAGAGAAAGGGCAACTCTACGTTGGAAAACT
GATTTGAAAGCAGAGATTAAAGAAAGTAACGACTAGGTAAGAGAGGATCTCTTTACCTGC
AACCAGGAGTCTCATAACTTGGAACCTTGCTGAAATACCCCATTTACTTTGTGTACTCAG
GTTGAAACTAGTACAAGTAGAGACAGAGAGGTGGCTTTATCAGGAGATAGGACGGCCATC
CCAAGATACCAGAAATGGCCTCTAGGTGGCATCTCTTTCCTCAGCTGTTTGAATTGTCTC
AAGACAAATGTTTACCTATGTAGGGGATTCAGAGTGATGAGAAATAGGGTGAAATAA
GATTTCAAATTTACACCTTTTAGTATTGCTTTTGATGTTGTAACCTACATATTACTTATTG
GAAAAATAAATAGAAAAAATATTGAATAAAAAACAATTTAAAGACTCTAGATAGGCTGGTG
TGGTGGCTCAGCCTGTAAATCCCAGCACTTTGGGAGGCCGAGCGGCAAAATCACCTCAGG
TCAGGAGTTCGAGACCGCCTGGCCAGCATGGTGAAACCCCTGTCTCTACTAAAAATACAA
AATTAGCTGGGTGTGACCGCATGCTTGTAAATCCCAGATACTCGGGAGGCTGAAGCAGG
AGAAATTGCTTGAACCCGGCAGCAGAGGTTGAGGTGAGCTGAGATCGCACCATTGCTCTT
CAGCCTGGGCAGCAAGAGCTAAACTCTGTCTCTAAATAAATAAATAAGACTTTAGGTAAG
TTATTTGTGATTACTATACATGAAATTGATCAGAAAGTCTAACAAATGATTGAGGTAATCC
TGTGTCCTTAAATCTCAAGCCTGGTTAACACAGTCTCTTGTCTGTGGGCCCTGTGGTA

ATCTCTGAACCCCTGGAATTCCTTCTAACAGGGAGTAGGATGCAAAAGGTGTGCGTTCTGCA
 ACAAGTAGAGGAAAAACAAGGGTGGGCTTCCAGATTCTCCAAGTGAACCTTAATACA
 CTGCCACCAAGAACTGCAGTTTCTTTTGTAGAGAAAAAGGTCCTCTATCACCCCT
 TGCAATTGGATGAGGTGACCTCAATGGGTTGCATGTTCTAGGCTCCCTGTGAGCTGGGA
 TGGGGCTTTGTATTTCTCTCTCCCTTATTCATATTTCTGATAAGGGTGAACTGTGGTGTT
 AAAAGCTTGGAATTTTGGCATGTGTGTGTTGTTCTGAATTCAGATATACTATAAACTGG
 GGTGTGAATATGATGGCCATCTGCATTTGATGACCTTGTTCCTCCCTGACITTA
 TCAAAGCACACATCACTCTAATGAAAAAGTCTGACTGTTGATTTCTTCAAATTCATGTTG
 GGGGAAGGTATATATTAATTATGTATACATTAATACATAATGTATGT

SEQ.ID.NO. 462 HG1015047N0 PLT00014333G02.contig

CACTTCTAAGGGGATTTTCACAAATCGGTTAAGCCCCCGTTAATCCTCACATGGTTATTGC
 GCGACCCCCATAGCCCCACGGTCCCAACGTTGTGATCTGAATCATCCGAAGCTCTTTCT
 AAAGGCAGGCTCCCCTGGCCGCCCTCTCCTAATCCGACTCGGTACTTCACGGGGCGTGCCT
 AGAATCTGCATTTTAGCAAGGGCCTCGGTGGACAGAAATTGCACGCTTGAGTTTGAGAAG
 TACTGCTTGGCCCTCTGCCCCACTTACCCTTACCTTCTTGAGATGAGGTAAAGCACTTGGT
 AGGCACTAGGCGGGTAGGTTTATCTTTGTCTCCCAAGTCTCTCCGAGAACTTGGAAGGC
 TGCCTAGCTTAGTGAGGGAATCCATTTCTGGCTTCAGACCAGTCTGGGTTTGAAACCTACA
 CATCCCACTAACTTAGGCAGGTACCGAAACTCCCTGAGCTTCATGTCTCATCAGTAAA
 ACGAGGTTAATCACACCTACACCTTAGCGCGATGGCTTAAAAAAAACCA
 AACAAAAAACATAGTCCATCAAGAGAAATACATTTTGCATACAGACCAGTACACTGC
 TATTTGTATGAGAGTGCATAAAGTTTCCGAAAAACAATGCTTACTACTTGTGACGCACTG

TGCGCACAGGGGCGGAGGACAGGGTGTGAATACTTGGAAGAAATCTCTGCAGTAGTTTTGA
 ATTATTTGATTGGTTTTTGCCCTTTTAAAGTCAACCTTATCCTTGTTTTGTGTTTCTAG
 GCTGTCTGACTCTTCTTACTAGACACAGCCAAAGAAAGGAAACCCTTTGTGTTTTTTTTTG
 TTGTTTTTTTTTGTGTTCTTCTTCTTTTTTTTTTTTTTCCCGAGATGGAGTCTTGCTC
 TGTACCCCAAGGTTGCAGTGAGCCGAGATTGCGCCACTGCACCTCAGCCTGGGGGACAGAG
 TGAGACTCTGTCTCAAAATAAATAATAA

SEQ.ID.NO. 463 HG1015088N0 PLT00014333G09.contig

ATGGAGCCGCAGAGAACCTAGAACCTGGGCCCTGAGGGCCGGCTGGAGGCAACAC
 AGGGCTGGTAAGAACGAATATGACCTCGTGTAAACTGGAAACGAGAAAGGAGGCAGAAT
 ATGGGTAGACAAGTCAGAACAGCTGCTCCCATAGCCGCCAGAACCCGTAGACTCTTCTGA
 ACTGTGGGACCCAAAGACTGCCGGGTACGCAGCGGGCACAGGACGGGCACTTGCGC
 TGCCGCCTACCGCCCTGGAGGGCGCTTCAGGCCCAGATCCCAGGGCAGCAGAGGTGCGA
 CAAGGCGTCCAGCAGCCGCGGCCCTGCAGATCCTTGGGGCACCTAGCCGGCATCCCCCTC
 CCTGACCCGACGCCCTTCTTTGTCTGAAGGATGGACTGCAAAGCTATGTCCGGCGGT
 GTGGACGAAACTCAGTAAAGGGGTTTCGGCAAGGGGAGACTCACCCGCCATCTCCGTAG
 GAAGGTGGCTGGGGCTGTGTCCGCGCAGCCCAAGTCCCGGCCCTCCTCCTCCCAAGGG
 CTTCCAGGTCTGACCGAGGGTCGCTGGGACTCGAGAACTGGAGGCTCTGACTGGAAACACT
 CCTGGGGCAGCTCTGGCTCAGCCTGACCCGGGCAAGGCCGTTTCAACCGGGGTGTCTAC
 ACTAGGGTCTCCCAAGGCAGGCGTCCGGGCGGCCAGGGTGGCAGGGCCGAGTC
 TCCGGAGGTGGCCGCTCTGAAGGCCGAATCTGGCCAGAAAATCTGGCCAGAAAAAGGGAC
 CACTGCGTCTTGCGGCCTCCTTTCCACGTTAACCGGGGCAGAGAACACCCAGGCCCTG

CTGTGGGTGCCTAACGCGCTGGGGCGCATAGGCCAAGCCAGTCCAAACCAGGGTGTCCAGT
CAGGGGCGGCTGAGGCAGAAAGGCTTTAGAGATCGAGGGACATGGGGGACCCCTGACGCC
AGGTCTCCTTCTCGCTTCCGGGTCCGCACTCTGGGTCCGCAGTCCCTCAACCCACCAGGA
GGCCAGAGATGCGGCTGCGGTGGATCCAGGAGCTGGGGCCAGATCCACCAAGATCAGG
GAAGGTTCTGAGGGCCCGCTGTCCCTCAGGGCCCTGAAGAAAATCCATTCTCCGCGTGCT
TGCGTTGGAGAACACAGTGGCCCCACCGCCCGCTGAGACCTCCACAGGGTCCCTGATC
CAGAACTTGACGACCGGACGCCATCTCATCTCTCCCCACAGAAACCCCTGTCAGCCCTGCCG
GGGACGCGGTGCCCGCCCTCCGGCATCCGGGAACACAGCGTCCCAGCCGGCCCTGCCG
CAGTGCCACACATTCTGTGTCCTCTCTGGCTGCGTCTGAGGGCCCTTAGCTCTGGCCAGGG
GCGCTGCAACGGGCTCAGCCTTCTAGGGGTGGCGGTGAAGGCCTACACCTGCCCCACAGGA
CGGCCAGGTCCACCAGCACCTGCAACAGACGGAGAACTGAAGGAGCATCTTGAAGCTGTG
GCTGTCTTGAAGGTGTGCGCCGCTTGTCCTCACGAGCACCATTTTGATTCTCGTGCCCTGT
GTGCATGGGTGCCAGCGTTGTCTGTAGGGTCCCTGCACGCTCAGAGTGGCCTTGGCCAG
GAAGTCCAGGGTTTCCGCAGGCAGAGTGTATGAACAAGGTGTTTATGGATTTTGTATTG
CCCATGTTTAATGAGCTTTACAACACTTCCATCAGAGAACGCTGAGTGCAGTGGTGCGA
TCTCGGCTCACTGCAACCTCGGACTCCCTGGTTCAAGTGTTTCTCCTCCTCAGCCTCCCCG
AGTAGCTGGGATTACAGGCACGTGCCACCACACCCGAGCTAATTTATTGTATTGTAGTAG
GGACAGAGTTTCACCATGTTAGCCAGGATGGTCTCAATCTCCTGACCTCATGATCCGCCC
GCCTCTGCCCTCCCAAAGTGTGGGATTACAGGCTTGAGCCACCCGCGCCGCCCTATATAC
TGTA TTCTTACAATAAAGTAAGACAGAGAAAAAGAAATGTTATTAAAGAAAGTCATAAAGA
AGAGAAAACATAATTTACTAATCATTAAAGTGGAAAGTAGATCATCACAAAAAGTCTTCATCCT
CATTGTCTCTCACATTACAGTAGACTGAGGAGGAGGAGAAAGCAGAAAGAGATTGGTCGCTGTC

TCATGGTGCCAAAGACAGAAAAGGTGGAGGCTGTGAAAAAGGGAGGTAAGAGGCAGACACA
CCGGTGTAATTTTATTGAAAAACGTCCACATATGAAAAACCGGCAGTTTAAAAATCTTGT
TGTTCAAGACTCTAGTATATTAAATA

SEQ.ID.NO. 464 HG1015049N0 PLT00014333H11.contig
TCTCAGAGGTGCAATTAAAGGGACCTTATGCAAACTGTGTGTTTTCTATAACCCATGGTTCC
CAGAAAGAGCTTGCCAAACCTAGAGCTCTGGGAACAAGTGGGGAGAAAATCTTAACATCA
TGCAAAAGGGCAACGGGTCCAGTAACATCTTTAACGTTATGGGCCTTAGTCAGGGCTGC
TTTGTCCTCCACTCTACACAGAAAGCCCTAAAAAGGGAAGGAGGAAGAACCATCATCTAC
CTTACCACCTCCTCTCTCCCTCAGCCCCCGCGTTACCGAGTAAAGGTGCCACAGAGGA
GACAAATATTTTCCCTGAGCCCTCTCTCCCAATATATTGAAAAAATACAAGGGATACAC
TACTGTTATGGGACCTGTCTTAGTCAAGTGGAAATTAGAAGGGAGCTCTTGACCTGCCCA
GTGATGCAAGATCAACAAGACAATCAGTAAGAAACGTACAAACATTAAAGAAGAGCCTATT
CAGAAAAGACAGTCTTCCAGTTACCTTATAAAGGCAAGGAAGAACTGGTAGTTTGAAATA
AAATAATTTCAAAAAGTGCTGTACCCCTGAGTATTCTGCTGTAATAAGGTAAGCACCTT
TCTCCTTTGTGATGCAATCATTTACATCTTCACCCCGCAACACCTGTCACATCCTTTT
AAATAATTTGTTTTTCCTTTCCCTAGGATGAAATCCCATAAAGTCAATTTTCTCGGGATT
GTGAAGTTCACCTGCACACAACCGCCTTATTTTACATCTCTTTCTTTTATTTCACCATC
TGGGGACTGGGCAGAAAGTCATGGTCAGAGGGGCACATATGTATTGCTTTGATGATGGGG
CAGAGACAAACAGGAAGAACTGGGAAATGAAAAATTGCTTTACTTCTTATATTCTGACTA
GAAAGGCAACCCCAAATGGGAAATCATAACAGGGATAGTTCAGTGGCAAAAAGACACCCCTTA
AAAATAGAAAGCTAGAAAGTTAAAGGAGGGGTGACATCACACTTTTAGAGTTGACATCAGAG

GAAGTTGCATCTCTTGGTGCCATGTTATACATGCACCCGTCGGGGCAGCAGAACTGATTAG
TGTGCCCAGTTGCACAAAAGGACTGTAAATGTGCATGGGCCCCAGTGTTAGCCATGTGACT
CTGTTAATGGATCCAGCTTGATCTTTACGCTCTTCAAATAAGCTAGTGCCCAAGCAGTTC
TTGTATTTTCAGCTTAGAGGGGTTCTGGAAAAGCTCTTGGTGATAGCATTTTGGGAAAC
GGAGAGAAAAACATTTCTATTAAAGTCTGCCTCAAAAAGTAGGAATTTACAGGTCAATGT
TAAATGCCCTGGCATGAGTTTATATATATTTTAAATAATTTACGGCTTATTTACCAAAG
AAAGCTGGTGAACACGTTTCACATCAGACAAAATAGGTATCAAAGCAAAAGGCATTAGAG
ATGAAGAGGATCACAGCACAGTGGGAAAAGGTTTCAGTTCAGATATAGGTACATATTATAT
AAAGCAAAAATACACAGAAACAAAAGGAAAATAATTCGATATTTAAAGATATTAGGTTAA
GTGAAATTTTATAAGTCTTGAAAACAATAATCAGAAAACAGGTGAGACACAGGAAAAATCATT
TTTAGCAACAAAGTACTAAGAGGGTGCAATGGACAAGGAAAACCTTTTCCCTCCCTAAGCCC
ATACCTCTAGTTGATAATCTGTTCCCATCCCCAGGTCCCAGCAAAACATTTCTAAGACCCCAA
GAACAGGTTTACTGGGAGGAATTCGAGGCACTGCAGTTCATGTGCTGAGTTAAATAAGCT
CTGGAATAAGAACTGAAGAATCACAATACCCAATCTAATTTAACCAAAAGTCTAATTAGA
GAAATTTACCAGAAAATTGAGAGACTTGCTAGACATAATTTCTGTAGGAAGTAGTATTTT
TCTCCTTTCAAATTTTAAATGTTTGGTATAATCAGCAAAATAGAAAGTAAATTATCAATT
TAAGCTTAACAGAGATATTTTGTTAAGTAATATCACAGTACCTTATAAAATAAAAAATATT
AATTATTTCCAGTTAGTCACACGTATTAATTACTAATAATGAGTAAATACAAAAATATACA
CTCAAAACAGAAAGAGCCAAACTGGCTCCCATAAATAAGTAGGTAATTTTCCCTAAGCCTTG
TACACAGTAATAAATAAATAATCTGCCACATCTTTACCATTCCTTAAATGACAAGCTGG
GTTTTGATGAATCCACTTTTTCAGTCAAAAATCATATTCATGATTTAAATTGTGCTGTA
TGAAATTGTCCCCTTTTATTTCTGTATATTCTCAGATGAGATGCATACATAATTGGTTCAAG

ACCTCTTGTTGCCCTCTTTTAAACATTGTCTTTGTATCACTGCAGCCCCCAACAACAGG
TTAAACACCGCCCAACT

[illegible]

TCCCCCTCCCTGAGGCTGGCAACCACAGTATGGTTCCCTATTTTCTATAATTTGTGCGCTT
 TAAGAAATGCTCTATAAATGGAATCATACACTATGCAACCTTTTCGAGACTGGCTTTTTCGG
 CTAAACGCAATTCCCTTGAGACCCATCCGGGTGGCAAGCATGTCCATAGGTGTTTCTTT
 TTAAATTACTTAGTGGTGTCCCGTCAATCTGGATGAACCAAGGTGTGTTCACACCGTTACACC
 AGGAAGGGCACTTGGATTATTTCCAGGTTTGGTTATTGCAAAATAGAGCTGCTGTGAACA
 TTCTTGACAGGTTTTTGTGTGAACATAGTTTTTCATTCTCTGGGATAAATGCCCAGGAG
 TGCAATTGCTGAGTCATATGGTTAGCACATGTCTAATTTTATGAGAAACTGCTAACTCTT
 TTCCAGAAATGGCTACACCAATTTTGCAATCCCAACCGCGGTGTGTGAGTGATTCAAGGACCT
 TCCCAGCCTGGGCAGCATTTGGTATTGCCACTATTTTTTTTTTTTATTATTAGCTCTT
 CTAATAAGTGTGATGATATCTCCTGATGGTTTAAATTTGCATTTGCTCTGATGGCTAATG
 ACATCAAGACCCTTTTATGTACTTATGGCCAATCTGTGTATCCTCCTCAGTGAAAGTCTG
 TTCAATGCTTTTGCCTACTTTTTAAATTGGATTGTTGGTTTTAACAACTGAGTTTTTCCTA
 ATTCTTTATATTTTCTAGATGCTGGCCTTACCACCTCGGTGATTATTTTCTACCACTTTAT
 CTTGTTTTTTCACCCCTTTCACAGAGTCCTTTGCAGAAAGTTTTACATTTTGTGAATTC
 AAGTTTATAAATTTTACCTTTTTTGGATTTTTGATTTTTGGTGTGAAGTCTAAGTCTTCAC
 CTAGCCTGAGGTTCCCGAAGATTTTCTCCTTTTTTCCCTACAAATTTTATAGTTTCACAT
 TTTATACTTAAGCCTATTTTCCATTTTGAGTTAATTTTGTATGAGGTGTGAGGTTTAGG
 TTGAGGCTTATTTGTTTTTGGTTTGGGTTTTTGCCCAACGAATGTCCAATAGCTCCGG
 GACTATTAGTTGAAAGGTGATCCTTCTT

SEQ.ID.NO. 466 HG1015052N0 PLT00014331I18.contig
 GCTGACCTTGAGAGAGCGGAGGAGGCGCAGGCGCAGCACAGGTGGCAATTGAAGCCGGAA

GAACATCTACCAAGAGCAGAGAAACCCAGGAAGAAAATTCTGCCTCTTTAATACGTTCCAA
TATGGACGTTTTCCATATAGATAACCTATCTATATAGATAGATGCTCTGGGATCTGACGGT
CCTGGACACCTGTATGGCTGTGTGCTCTTTGCCTAGCCTGCGGTTCACTTTTGCT
CTGGCCACCACCTCCCTCATGTACAAACCGCTCTCTGCTCTGCCAGTCTTGGCCCCCG
TCAGGCAGCGGTTCACTCCCTCACCCAGGGAAGTCTAGATAGATAGATATAGATAGATATC
GATCGATAGATATAGATATATATAGATAGATAGATAGATAGATATATGGATAGATATA
GATATTTGGATAGATATGTAGACGTCTATCTATAAATATATCGATATATCGGGGTTCTCA
GCCATCCCAAAGCTTGCACAGGGCCAGTTGCTTTATGTCAGGAAAAGGGACACAGAC
TGAGTGTATGAGGGTTTTTTTTTCTATAGATGTATAGATAATTTATATCAGATGCAATA
AGTGAAACTCTCATACAACTTTTAATTTTGTGTTTTTCCAAACTCTATTCCATTCAACT
TTTTCTGCCTTTTTTATTGTATTATAATGAGACATCGTACTGGAGCAGGCTTGAAATGGA
TATACAGAGCACTTTC AACCCAGAAATTTGTCTCACGACTTGGGGTCTTACTTAAAGTTG
TAGCTGTTAGTAATTTTCATCCATGAATTCACCTTGTGCCATAGTGTGATGTTTTGTTATG
AAGTTGGAAATTGGGAAGGGAAACTAGGAAGTTTTTGGGAGAAAGTTAAGAAAGGTTGAGTAGG
CAGGCGTTGTTATAGATATTAAATTTCACCCTGCTTATCAAAGTACCTGGAGCGTATGGT
ATCCAAGCTGTAAATGGCCCAAAGGAATCGGGGTTCTCAGCCATCCAAAAGAACATGCA
CAGGGCCAGTTATTTATGTCAGGAAAGGGGACACAGACTGAGTGTATGAGGGTTTTTTT
TTTTGTTTTTTTTTTAGTTTTAAACTTGATTTTTGTGAGGGGACGAAACAGCGTGACAAATT
AAAAGAAGATCTGGGAATCGTAAGTACTGACAGAAATACCTGGAGGGAAGGTGGGACTGGA
GAATAGTTTGTGAACAGATCCTGCTGAGGATGAATCCAGAACTTCTGAGGGTTCCGC
TTCTGCTGTAGGGACAAAGGTTGTAAATAGACAAGCCACAGGCACCTGC AAAATGGTAAG
TGTGTTAGAGAGGGGCTGGAATTGCCTGTTGGTGTATCTGTGGGAAGCAGCTGCTATCCT

TGAGGAAAGAAAGATGGAGAAAGGTAAAGTGGCTATCAACCTGTCTCAGAGACAACACATTCTG
 TGAAAGTTGTAAGTCTTTTCATTTCTGACACTGTGGTTCTTAATAATCTGTCCTAAACCTGGCA
 GTCATTGAGGTTTCGCAAAAGTATCTCCTTCTATGTGACAGTAAAACTGGAATTACTATCAG
 AATGGAAATCCATTCCCTGCTAGAGTAAACCATGAAAAGAAATCGCAGTTTGTGATTATCTG
 CTTATCATTTTAAATAGATCACAGAGTGCAGTGTCCCTTATTTGTGTATGGCAAGTTGA
 TGAATAACTCTTGATTAAATCTTTTAAATGTTCTGTTTCACACAGTTCCTCTCAATTACAT
 GTGAAAAATCTGTTATATAAGTAATGCTATTTAAATGATTAAAAATGGATATATGTTTTAAG
 GTAGTGCCACCCAGTCTTTTTTTTTTGTGAGAATAATAATTATCTATTGTTGCCTGC
 CTTGAAGTACAGTAAATTTCTCTGTTTCCATTATTTAAAAATAGGTGAGTCTTTTGACAAA
 GATACAAAGAAAGGGACAGTTATGAAAGGGAAGAAAAGATAGATGGAGATGAAAAAGAAA
 TTCAGTAGTTTCAGCAAGGGAATTAGGAGCATAAATCTTTGAGAAAAGTCTTGGTTGCTTA
 CCATTCGGGCTTTGTGATGAGTTGAAATGTCCTGAATGAGAACTGTTCTCAAATGCTTTC
 AAGACCAACAAAAGTACGATACCCACACTGACCTGGAACCTCTTTAAACCATTAAGATAT
 ATCACTTTTTCGCAGTTCAGTCCATAATTTTTTGCAAAATGTTCTGCTTAGTTTTCTGCT
 AAGAAATTACTGTTTCTGAACAAGGTTAAGTGTGCGTATGCTTTCATTGTAGTTTTTTTTC
 CCCTAACATATACTGTTTTCCCTAATTAGA

SEQ.ID.NO. 467 HG1015054N0 PLT0001433122.contig
 GAGGTTGTCACTAGAATGTGAAGACAGCCACACAGATATTGCACAGACTATTACAGAT
 CGTTTGGTTTACATTGAGAGTCATTGCTCTACTTTTGTGCGGTAGGAAAATGAGATTTC
 GCAATTCCTTTTTCGATTTTTATTTTTATTATGAGTCTTCTCCTTATCAGCGGACAGAG
 ACCAGTTAATTGACCATGAGAGAAAACCTGCGCAACACAATTGCCCTTCAGAGGAGATG

TATGCCTCTCCATTACGAGTACCCCTTCCCTGAGATCTCTCTAGCTAACTTTACTGGAT
CTATCAGAAAGAGAGAGTGAGGAAAGACACCCAGCCACACAAAGAAGAACTTCATGA
TGCCAAACAGCGTGATTGCTTAGAAGTTCCTACACAAAAAAGGATCATTTGAAAAGCACCT
GGAATGGTTTATTAGCTTCACAGGATTTTATTCTTCTTGGCTTCTATTGAGGGGAAAAAT
AACATAAAATTCAAAAAGGATTCCAAATCTGAAGCCCCAAATCGTTTGCCCTACATAACAAAAAT
ATCTCATCTTTTCCCTGCACATTATTATTCTTTTATGGTTAAAAAGAAAAATACCTTTTA
GTGTTTAGAACTCTCTCATGGTAAAAAGTGCAAGAAATTTAAAAATGTTGCTTTCATATTC
CTATAATTCTCCAAAAAGTATTAAATTCGTATATGTTTGAGTGATTTTCTAAAAAACTGCTC
AACCTGAAATCAATTGCATTGACCAATTTGGCTTCGCACAATAGGGAGAAAAATAATTGGTT
CATTGATTATATAGAGAGAAAAGACTAAGAAAAAGCTATTAAATTGCTACCAATTTTATGATA
AGCTTTAAGGTTTATGAAAGTATGTTTATTTTATTTAATGAGTAAATGTCCATTTGAAAGTTG
AAAGAAAAACATGAAATCCTAATTGTAGTTCATTTTATGTTCAAATGAACCATTTGTTTTT
GTTTTTGTGTTTGAAACAGAGTCTCACTCTGTGCCCCAAGGTGGAGAGAAAGTGCACGCTT
TTGTCTCACTGCAACCTCCACCTCCCGAGTTCAAGTGATTCTCGTGCCCTCAACCTCCCCAA
TTATAGGCTGGGATTACAGGTGTGCACCACTACACCCAGCTAATTCTTGATTTTGTGTA
GAGATGAGGTTTTACCCCTGTTGCCCAAGGCTGGTCTTGAACTCAGGCTGGAACCATTCATT
TTTTAACCTTTCTCATCATGTAATTATAGGAACCCAACGTTTGATTTCCCTTTGAAGTTTT
GTTATGTCCTTTATTATTTGTATGGATAAATTTCTTTAAAGTCTTACTTAAAGTTGACA
TCTAAAAATACAGTTATGCCAATGAAGTCCCACTCAGGGTGATATCTGTATCTAAAAAGATG
AGTGCTCATCATCCTATTAGGCTTTGTCTTGGTGGTTCATCCTGAGATGCTGAGACAT
GGAAATAAAAAATCAGAAAGGAATTTAGGGATATGATTACTCAAAAAAGAAACTATCCTGT
CTAAATTTGAAATTGTGTGATAAATACTAGGTGTTCCCCAGATGCTAAGATGTTCTTAAATTTG

TATTTATTGAAGGATTGTTAGCTTAGTGCCACAAAATTTTCTTACTTTATGTTAATTCC
 AGATAAGAAATTTACAAAGTTTATATCTTTTTTTTTTCTTTTTTTTAAAGATGAGATCTGGCT
 CTATCACCCAGGCTAAAGTGCAGTGGCATGATCTAGGCTAACTCCCTGGCTCAAGCGATC
 CTTCCACCTCAGCCTCCCAAGTACCTGGGACTACAGGCACTCACGGCCACACCTGACTAA
 TTTTGTATTTTTGTAGAGATGGAGTATCGCCATGTTGCCCAGGTTGGTCTCAAACTC
 AGGCTGGTGAGCTCAAGTGATCCGCCCTCCTTGGCCTCCCAAATACTGGGATTACAGGCA
 TGTGCCACCATGCTGGGCCACAAAGTTCATATCTGGAGTAGAAGTTTACTTTGTAAATAT
 TATAAAGTAGAAGAAACCATAAACCATTTTGCTAAATGAAGGTTGGGGTTAATATAAA
 TGTAAATTTTAAATAGAAAATCTGACAAACACTGTGAGTTTGTCTTCTGTCAAAAGCTTAT
 TAAAAGTGTCTTTGCGGATGAATGGTACTTTCCACAAGTGCAATTTGAGTAGAAGCATAAC
 CTATTCTCAGTTATATTTATGTTTTAAACATGTACTGGTTTGTATATTTTGTACTGAAAA
 AGAAAACACTTTATAGTCAAGATACATCTCATTCAATACAAAGTCTAAACTCTTTCAAATA
 CAAATTCCGATATTACACAGAAAAAGTTACAAATCAGTTTTTACTATTGTAAAGTAATGAAA
 TGGTTATACATTTCTTAATTGT

SEQ.ID.NO. 468 HG1015056N0 PLT00014333J01.contig
 GGGCCTCACGGGAGCTGGATGCTCGGAGCAGCAAGCGGCCCTTGGGGCCCTTCCGGTC
 CCGGGGCTGACGCCCTGGCGGCGTCTCCGGGAACCCGCGCCCTTTCAAAGCGCCGCTGGCTC
 CTGTTTCACAGATGGAGACCCCTGAAGCCCCCAACAAGCAGCCTCCCCAGCACCTCCCCAGTG
 ACTGAAAAAGCTGGCACTGGGCGCCCGTCTCACTCTCCACCCAGCTCCAGGCTGCCCCCA
 CGGGACCAGCACCTTTGGGGAGAGACGAGACCTGAACAAACGCTGGGGCTGGATCCCCAG
 GGTGGGTCAGAACAGCCGTTGCCCTGAAGCTGGCCCCGACAGATGCGTGGGGGCGGTGC

CACCGGAGCTCCTTCCAGCTTGCGTCTGGGGCCACCATGTCCCTTCCTCAGGGCACTGGA
 TCCATCTCTTCAAGCTCAGTGGTCCAGGCCATGGTGGCAGAAATCACCCCCCAGGGTGTGA
 CGCCTCGTCCCTGTCAATTCCACACAGCTGCA CCCGGAAGAGCTGTGGCCGGCGCCGCT
 CTGACCTGTGTGCGGTCCCCTGAATGGGTCTCAGGTGTCTTCAGAAATGGTGAGGCCGCCT
 CATTTCCTCACCAACAACCTCCTGAGGACGTTGCCCTGCAAATGGGGGGAATTCCTAAT
 TTCTACCATTCAGTTAAACGAGGGATCCTTAAGCCCCCAAAGCTGGGAAAGCTGTCCGTTA
 ATTTAGTGGCGTGGGCCGTCTCAGAACAGCAGGGGTTCCCAGGAGTCCAGCGGACGGGG
 GAGGGGATGAGTGGCCGTGTCCCCCTGGCAGAGAAAGGCCTTGTCTGAAGGCTACGCCCGC
 CTCGGGTACAGGGACACCTCCTTGCTCATCTGGCAGCAGCAACAGCAGAAAGTTGGAGTCG
 GTGCCACCTGGGACGTACCTGAGCAGGAGCCGAAAGCATGTGGTACTCACAGTATGGAAAT
 GAGGCCATCTTGGTCCGAGACAAACAAGCTCGAGGTCTCTAGGGACACAGGGCAGTCC
 AAGTTTTGCACAAATTATGTAATTCCGATGTGAGCACCTGAACCCAGGACCACACTTTGAG
 GAAACACAGACCTGAGTTTCATTTCAGTTGTCTTGTGATTTCCAATCCTTGCTGGAAGATT
 TTGAATAATGAACGTGAAGATCAAACTGTGGTGTAGAACCCCAAGAAAGCGCAGAAATTGGC
 CCCCGAACCCCTCAGCACAGCCGCGGTGTGACTGGTGCACAGGACACTTCCCCCTCTAGAG
 TCCCCCTCCGTGGTCAGCTGTTGTGACATTTGAATTCTCCAGGTGCTGTAGGGACAGCAT
 AGGCCCCGGGCACTGCTAGATGCTTGTTTAAACCAGGCCAAGAGTAACCCCTCCCATCCCC
 ATTACAAACCACTTTTCAAGTTTTTTAGGTTTTTAGTTTTTTTGTATTTTAGCATAAATGGC
 GGAGTAGCAGCAGCATCTCACCCAGGAGCTTTGTGCGCATTAACATTTTGTATTTTGTCTTTT
 TGAGACGGGGTCTCACTCTGTTACCCAGGCTGGAATACAGTGGCACAATCACAGCTCACT
 GTAGGCTCTATCTCCTGGGCTCAGGTGATCCTCCCAAACTCAGCCTCCCCAGGTAGCTGTG
 GCTCCAGGCGAACTACTGTATATAGAGACAGGGTTTCCCTATGTTGCCCAGGCTGGTCTG

GTCTTATAGTCCTGGGCTCAAGCCTCCCAGAGTGCTGGGATTACAGGCATGGGTCAACCGC
 ACCAAGCCCGCATTTATATTTTGAGTTACCAAAATGATCATCATCTGTGAAGCATCATGG
 TAGCTTCCAAAGCATTTGCGGGCCCACTGGCCAGTGCAATCCAAACACATTTCAGGAAACCCG
 AGAGTTGTAAAGTTGCACATGTGTGCCCTACAGTGACTTCCATGTAGTAACAGGGCTGAGC
 AAGGCAGGGCTCTCCACCAGACTTAAAGCATTCAACCACTCTGCCCCCAAGGCCGACACTGA
 CATGTGTTCACTGTCGCTGCCACAGCTGGCCACCTGTGTTCAGTCTCCATGAAGGATTCA
 CCTTCTGCTCTCACTGACATCTCCTTGGCTACACAAATGTGCAGGTGTCTCTTGTCAAAG
 CTAATTTTCTTTTCTA

SEQ.ID.NO. 469 HG1015058N0 PLT00014333J13.contig
 CATTTGGAGGAGGGAATGCGACTCTGAGCGGGCCAAGCCCTAGTGGGTCAGAGTAGGATT
 CTGATGGAATTTGCGGTCTCACCGCCTACTAGGGAGGAAAAGCAGTTTTTCAGGAGTCCCCG
 CACCGCCGCCAGCCGTGGGCTGGGACTACAAAGTCACGGTCTGTGGAGTCCAGGAGTGGGG
 GCGGGCGGGAGGCAGGCTGGGAGCTGTAGTAGTGTACCCGCCACGTTCCGCAGCAGGTAGA
 TGCTCATCTTGCTAACGGATTGGTGAGAGTGGACAGCAATTAGGCTTTGGCACCTTGGAG
 CGTGGCTCCAGCTAAGGGAAAGCGAAACGCTTCCTGGGTGAGAGGACTTTTATTCTTTGCT
 CATCCTCTTAAACCTGGCGAGTGAGCGCCTGAGGTTCCGGAGTGCCTTCTTCACAGACGAC
 GCGAATGTGGGGTCTTCGTGTTTGCCGAACTGAAGGAAATTTGCCCTTCCTGCCCTTAGGT
 TCGCAGCAGTTCATAACTGAATAAGAGACTTTCCTCCATAGAAATGGACTTTATAAATCTA
 TTAAATTTTTTTTTTTTTTTTTTTTAGAGCTAAGGTCTTGCTCTATCTCCCAATGCCAAAG
 TTGCAGTGAGCTGTGATCACACCATTGCAACCCCAACCTGGGCGACGGGAGTGAGGGCCCTG
 TCTCAAAAAAAGAAAGAAATAGGTGTAAACAGGATATAGTTTATTACTTCT

TCAAAAAAGGAAAAATATTTCTGTTAAAAATAATCCAACTAATACATCCACTCCAGAAATA
 TTGCACCTGGCTCAGTAGTTGAAAAATTTGTTAAAAGGTATTTACTGTTAATAGCTAGTTA
 CGATTAGTCTTAGTTTTCAAATTAGCAAGCATTTTATGGTTTTTAAAAATTGTTTTTAATTTC
 CATTTAATATGCTCTACCAGTTACTTTTTAAAGTATACTTTTAAAAAATAGTTTATTTTAA
 ACTTTAGGGCATATCCCTTTTTTTTTTTTTTCAAGAAATTAACATAATGTAATTTATCAA
 GGCAATCATCAACTGCCCTTCCCTTGTCATTTTAAATAGTTCCATTATTTCAAAATTC
 TGGTGAAATATTTATTTATATCTTTCTTCAGTAGTAAGTTGTCTCCTAGAGATAATTTTCAT
 TTCTTTTTCTAGCTGAATGTTTGGTATTTAAATATTTCTTCAGTAGATTTTCAAATTTGCTAT
 CTTTCTGTAGTTCTCACTGAGTGATCTTTCTGTGTGTGTTACATGAATACTATTTTAAAGT
 ATATTCATAATATTTCACAAATCCAGTATCATGTGTCAACAGATGAGTCTAGGCAGTGG
 TTTTATTTCTAATGAGCAATTTGTGAGACATACTATTAATCATGCAATACCTTTAACAGT
 TGAACTTTTTTAAATTGTGGGTTTGTGTGCGAGGGAGATATTTATGATTTCTACTAAAGC
 CCTCAAAGCTGCTCAGACAGCCCTTTTGTCCTATTTTGGTCAGTCCCATGGCTTTATCA
 CTCGCTTCAAGCCCTACACCCCTGCTTCTTCATTTCTTAGAAAAATGGAGTTATCAAGTTTT
 AACTCAGTTTCTCCTTTCCAAAGTGAAATTGGGAATCACAAATTTACCCCTATAACTTAC
 TCATTTCTCTCCACTATTAGGATTCTGTTGACCCCTCTCTACCCTTTTTATGTCTCTTCCTA
 CTCTTGCTGTACAAATATCTCCTCTCCTTTATTGTTCAGTCTTTTCCCCCTTCAGTGACTCT
 CCTTATCCTACAAAAATGCTCAAGACTCCCAAACTACTGCCCCATCTCATGCAACCCACA
 ACAGTCTGACTACTGGCACATATCTCCTAAAGGACTTCCTACTTGCCCAAATCTCATGACC
 TCTTAGTGTACATCCTTCTTAACCCCTTGTGCAGCATTTGACCCCTTTTCAGTGAATACCAT
 TCCTTCTTGAAATTCTCCCTTTCCCTTAGTTTGTTTTAAATTGAAGTGAAATTTATATAATA
 TGTAAATTAATCATTTTAAAGTTAACAAATTCAGTGGCATTTAGTACATTCACAATATTGTG

CAACTACCACATCTATCTAGTTCTAAGACATTTTCGTGTCCCCAAAAGAAAGCCCTATTCC
 CCATTAAACAGTTACTCCCTACTCCTTCCTCCCCTAACATCCTGGCAAATCACCAAATCTGC
 ATTCTGTCTATATGGATTTACCTATTCTGTATATTTTCATATAAAAATGGAATCATATATA
 ATAGTATGTGGCCTCTTGTGTCTGGCTTCTTTCACCTTGCAATAATGTTTTGTTCAAATCT
 TAATAATTTTCATATAATTTGCCCTCAGATTTTCAGTAAGAAAACAGTATGGATAGAGTTGAGAT
 GCCTATAGAGTCCTTCTTCTCTCCCCAGAAATAACCAATTATCTTGTATTTAGAAATTTT
 TATTGTTCCAGTGCA TGTATAAATTTATGCTGAGTAGTATCTTGCCTTTAAAAAATTCT
 GGTAAAATCACATAACAAAATTTACTATCTTAACCAATTTTAAAGTGTTCATTAGTGTAA
 ATATAATTTGCATTTGTTGTGCAACAGATCTCCAGAAACTTTTCACTCTTGCAAAACTGAATA
 TGTATA

SEQ.ID.NO. 470 HG1015060N0 PLT0001433J15.contig
 TGGTGATACAAATGTGGGACTGATTAGTGAGGACAGAGACAAAGAGGCTATGCAAGCAGTTA
 GGAGGCTGCTGCATTTATTACTTTGGCTGAGAAAAGGTGAGGAGGCTGCAGAAAGAAAAGTTG
 GAAGCTAGAAGAAGTTGGTTTATGAGGTTTAAACGAAAGAAAGACGTCTCCATAACCTAAAA
 GGGCAAGGTGAAGCAGCAAGTGCTGATGGAGAAAGTTGTAGCAGGTAATCCAGAAAGATCTA
 ACTAAGGGGCACACCCAGCTCCAAGATCACCTTCCCAGGCCCTCTCTTGCTATGACAGATG
 CCATGTGACACAGTTCTAGCCCAATGAGAAAGTAACCCAGAACACCCCTGGGTGCAGTGGCTTA
 TGCTGTAATCCCAGCACTTTGGAAGGCTAACGTGGGAGGATTGCTCGAGACCCAGCCTGGG
 CAACATAGTGAGAGCCTGTTTCTACAAAACATACAAAATTAGCTAGGCGTGGGGGTGCA
 TGACTGTAGTCCCAGCTACTAGGGAGGCAGAGATGTGAGGATTGCTTGAGTCCAGAAAGGT
 GGAGGCTGCAATGAGGTGGGATCATGCCACTGCACTCCAGCCTGGGAGACAGAGTTTTTGC

TTGAGACCTTGTCTCAAAACAAAATAAAACAAACCACCTCCTACTTACTGAACATACATTA
TATATGTGTCACCTCAATCATACTTTATCACATTTTAATATCTGCCCCCATTTTTTAAAGT
AGCTATTTACTGATAAAGACACTGAGGATCACAGTGCCTGTAAATTAGAGCTGAGATGAA
ATTAAGGTGATATTTAGTTCCAAGCTCCTTGCTATTTCCAAACATTACTGATCACCTCTCA
ACTGCCAAACTCTTAATAAGTGAAAAAAGCTACTTTTCATTATCATTTGAAGTTGCCCAATTA
AGCGCTCAACAAATAAGTAGCTCTAAACCTGCAATACCAAATCTCTTCAGGTGTTTCCAG
ATATAAATGATGGACACACAAGTCTCTACCACCATTACAGATAACAAAAGGCATAGATAT
GCAGAAATAAGTATCTTTAGCAGCCTATGGGTAAACCCCTGAGAGGAAAATTTGTTTTCCTT
TACTAAAAGAGAGGGAGAAAGGCAGTTGAGAGAAAGACCTTCTGAAATAGGAGGAGCTGTATG
ATTATCCTTCTGTGTTTCGTTGTGAAAAATATGTTGGTACTCTGATCACTTCGGTCACITT
CCTAGGGGAGCTTTTCTTAGTTTTTCAAGAAACATGAAAAACATAATTGATTTAAAGGG
CAGAAAGGTTATTATATGGAACCCACAGAGTGTGACATTTTATTCCTGTTCTGCAAAATTTCT
TTATCAAGAAATAATACTTTGAGATAAATGTTATAATAATAAGTCCTAGCTTTCTTAATA
TTTCTAAAACAAAGATGACCAACTTACACCAAAAGTACCAAAAGTAGAAATTAAGGTTATAA
CTACCACGGTAAGTTACTGCCATTAAAAAGTACTGCTGTTCTAAGTTTCTGGCTGGATAA
CAGACTACCATAAAGGATTATAAAGATGAAAATGGCTTATTTTTTGTGTTCTAACATGGAAAA
TATTTTCAGGTAATTTTTATTATAAATAGAATAATGACAAAGATTCAACAAAAGAGTAGGT
TTTAAAGCATTCCTGAAACCCCTCTGTAACATCTAAGCTTGGTGAATCTTTAACAGTCCAGAA
GCTTAGTTTCTATAATTGCCCATACCTAAAAATGTAAATATCAGTAATTTCTTTGTCTTTGA
TTTTTTTTTATATCCATTAGTAGACCCCTTTTGACCTCAAGTGTAAAGAGGATTCCTTGGTAGAT
TAAAAACATAAATACTGATATTAATAGTTTTTCTTCAGATATTAATAAACACATTGAAACA
GATTAAACAAAGGGATATCTTTTATAAATTTTCAGCATGACTCCATATTGTCAAGGTGACTG

AAATTCAGTAGGTACAATAACGTTGCTTATTGTTATTGTTACTGATATTATTTTCCGG
 TGAAAGTGCAGCCGTGGATCTTCTGATGCACAGTGATGGTACCTTGCTCTATTTTCCATATA
 TCACATTTATCCAATTACCCAGTCATCTGTCTTACTCTTTTCTTATCACCAACATAACA
 ATTTTAGCCCTTCATAATGGCTGGAAATTGGCAACACAGCAATGAGGGACTACCACATATCA
 AAACCTATCACAAACCACCTTTGGTGTAATTCGTATTCCTGACTGAGAAACAGAAACAATCA
 CAATCCAAATGTTTCAGAAATAATGCATTAATATGAGCAGTGGCAGAAATCTATTTTCAC
 TTTTTCTGATACATTTTCTCTCATATTTTCATAGTTAGAAAGAAATCAAGAACATTACAA
 AAAGATGTTTTTGGCTTTCACAGTGACACCTTAAATTTTTTTCAGTTTGGTTTTAGTCTA
 TTAATTGATGGTGGAATGAATAATTTTGCCATCAGCAACTGTAGTATGAAAAATGCAATTA
 AATATACCTTCCAAGAAATTTCTATTGAATATAAAATTTGAACCATTTTAAATATATTTT
 TTCTAATTTTCAAGAAGACATTGGAGATTAAATAATATATGCCCTCTTATTTTATATAAAGC
 AATTAAACAAATTTAAATTTGGGTAAACCTTTATTTATGTATTTTCTGAAACAGTTAAA
 CACATTACATATCTAGCTCTGAAAAGTATAGCTGAGTTATTGGGTGTTTTTGTAAATGAA
 GACAAAAAGAGGAAGGAAAGATTGGAATCAAGTGTGTATAGCATTTTATATCCTGATTTTC
 TTATAAATAATTTTGCATGGTCACTAAATAATTTCTTTTGAAGATGTAATTTATAATGGTTACA
 GGGCATTGCATTTACTGTATGTGTGTGAATTTATTCAGTTTCATCTTTTATTATTGAGTTT
 TGTATTTTTTCTATTATAAATA

SEQ.ID.NO. 471 HG1015062N0 PLT00014333J17.contig
 GGGCTGCTTTTGCTTCTAACGTAAAGTCAAAATAGCCAGCATGGAAATAACCTAGGAAAAAA
 GTGGGATATGCAAAAGCAAAACCTCAAAAAAAAAAAAAAAAAAGACAAAAGCTGGTGGGCT
 GGCCACGGGTGCACAAATGAAGTCTCCAGTGGCTTCATCATATAACAAATCCCCATCTGTAGG

AGGTAAGCCAAACAAGGTGTACTGCACAAGAAAGAAATGCAGAGCTTTCAGTCTGGCC
AAGAAGACAGGCTAGGTTCAAGCAGTACTCGAAGATTGCTTGTTAACCCACGATAGAGG
GTAAGAGTATACCCCTTGGGGTCAGACTCCCTGGTTTCAACTCCTGGGTAACTTAGAAAGT
GTTACCTAAATTTTCAACGTGTTGGTAGACTTTTTCATCTGTAAATAGGAAGCATAAG
AATTTTAGATGGCATCCCTACCCAAATCTACATAATTTTGGCGAACATCATTCCTCCTACT
CTCTCTACTTGCTAAGGAATGACATAGATTGCCAGAAAACAACTACCGAAGCTAAATGA
GAGACATGGAAACAGTTCCCTCCCTCACAGCCCTCAGAAAGAAACCAACTCTACTGATGTCT
TAATTTCACACTTCCAGCTTCCAGAACTGTGAGACAATACCTTTCTGTGGTTTAAAGCCAC
CTAGTCTGTATTACTTTGTATAGCAGTCTTAGCAAACTAGTATGGAAAGGAGAGCACAG
ATAAAAGATGTGTTATCAATCAAGTCACCAAGTGTGAGCAAAATAGATCCTAATCCCCACTGG
AGAACTCTGGCAGACAGTGTAGAAATGTACAACTCAAAATTATCCTGCCAGACAGACCCCA
ACAGTAATGGGTAAAGATTTGAATAGATGTTTTTCCCCAAGAAAGATATACAAATGGCCAGT
AAGTACAAGAAAAGATGTTCAACATCATTCATCATTAGGGAACACAAATCAAAAACCCACA
ATGAGATATGTACACCCACCAGGATAACCACAAATTAAAAAGACGATAATAACAAATGTT
AGTGAGAAATGTGAAGGAATTAGAACTCTCATTCATTGCTGTATGAATGAAAACTGGCAT
AACTGCCTTGGAAAAACAGTTTTTCCAGTTCTTAAAAAGTTAAACATAGAGTTATCCTATGA
CCCAGAAACTCCACTCCCAGTTGTACATACCCCAAGAAGATTGAAAAAAATATTCGCATG
AAACTCGTACATGAATATTCACAGCAGCATTATTCAGAATAGCCAAAAGCGGAAAAAAAT
CCAAAGTGTCCATCAACTGATGAGTGGTTAAACAAAATGTGGTATATCCATGCAATGAAAT
AGATTTCAGCAATAAAAAGAAATAGATTATTGATATATGTACAAACATGAATAAATCTTG
AAAACATTATGTTAAGTGACAGAAGCCAGTCATAAAAGCCCCACATATTATATGACTCCAT
ACAAATATGTTTTATATAAAATTTATATAAAATTTCCAGATAGGCAATCCCATAGAGATG

GAAATAGATTAAATGTTTGCTAGCAGCTGGTGGGAGGGTGAATGGGGAATTACTGCTAA
TAAGTACAGAAATTTTGTGGAGTGATGAAAAATGAACTGAAATTAACAGTGGTGATGG
TTGCGCAGCTATTTTGTAGTTAATATACTAAAAACCACTGAATTGTACACCTTAAAAAGTAT
GAGCTTTATGGTATGTGAATTATATC

SEQ.ID.NO. 472 HG1015063N0 PLT0001433J23.contig

CCTGGCCTTTGACGTGTCTGTTCTCTCCTCTCTCACTCCACCACCACTAGCTTCCAGGGA
TCTCCAGCCTGTATAGAACTTCTCACTACGCTCTGCTCTCGCCCTTTTAGCAGTCTCCT
TATGCTGCTGACTCCACCTCCAGGCAATCATTTCTAAATTTAAATTTCTTTACAAAGTATTCT
CCCTCCTGGAGTGTGGGTCCACCACGGGGGCAATCAAGGCACATCACCCACAGCTTTC
CAGACACCCACAGCCTTCTGTCTGTGTGAGCGTGGCTCCTTCAGCCTTCTTCCTTCT
TCGCCGTGGGGAGTCAGACATTCGCCAACCTCGTGTCTCAATCCCTGCCCTCTGTCCAGGC
CTGCTAAGCCCATAGACATATCTGCATGTGTGACAGAAATATGTGCAAGCTTTCGACAGGAT
CCCCGTTTCTCCCCATGGAGCTTTTGGCAAAATGAAGGGTGTGGTTCAGCAGCTGGTCTGA
TGTCCCCTTAAGCTCCCCACACCAGCACACCAGGACTGTCTCTGTCTCCACATTCACAGCA
TATCTGGGTCTCAGCTTTTGACACAAATCAGACACTGTTGCTGAAGTTTATGTTTATATG
ACAGAAAGAACTGTCACCCCCTCCCCTGCCCCCCCCCCCCCCACCCACCTGTTATAAGAAAAG
GCTCCAAGGTCTTGAGACCACACGACGCGGGAATGGGAAGCTGACATTAAAGGGGAATGAG
GTCTTTCTGTGATTCAGCAGCCTCAGGAGGACAACGTCAGAAACCAAAAAGCTCCTCTGC
CCAGGGAATGGGACCCGCCGACTTCAAGGGCAGAACAGGTGCCCTGGCTAGTGGGGAGTT
CCCCAATGTGGATTTAACTGTAGTGTCTTCAGAGAACATAACCAACATTTGTCTTTAGG
AGATGTGTTTGTGTGATGATCAGCTATGAAAAATACCCAGCGGCTGCAGCCCTCCCG

GCAGATGCTGAAAAAGGAAACCCGACTGGGAGATGCCGCCACCCCTTATCCCCAGAGAGCT
 GCGTTGTTTCCTCTTGCCCTTGGCCTTGGATCCGGCCATAGCAAGCATCCCTGGGCAOCT
 GGCTGAAAGCTGGAAACTCGGGATAAGAATGGAAGCCCTGGAAGGGAAGAAACCGAGGA
 GTGAGCTTGCCCTGAGGACAGCACATTGCACTCTCACCTCCTTTGCAAGTTAAGTGTGAAA
 TAGTGCCCAATAGAGTGTAAAAGAAAAGAAAGAGATGATTCACCTCCCAAGATTCCCCA
 CTGCTGTGTCCTTCTATCAACTTGATACAAAAAACACAAATGGCCCCGGGAACACAGTG
 TTGAAGATGAGAGAGCTAAAAGATGGAAAGAGACTGGAGCCCTGAGTCACCCACCTGGAGG
 AAAACTACCCACCCATCAGGAATTTCTACTACACATTTGGGAGTTTGTGGGTTACAGCA
 GCTAATTTAACGAAAACACCCAGAGACAAAGAGCTATTACTGCACAAATTGGTTAAGTTAC
 CTTTTGTATGTTTGGGTCCTCTCTATTTCGATGAGACTGCGCGTTCCITTTGAAGCAAAATC
 TGTTTTAAAATTCTTATATCCCCAGAGGCCCTAGCAGAGAGTATGCACAGAGTGAGCCCTTC
 AGTAAATGCTGAGTAAATAAGCTCCTAATGGATGAAACTGTGTCTTCGAAAGTTTGTCTG
 ACCAGTACTTAGTATAGGATTTTGTGCTGTGTGTCACCTTTTCATAAAGGTTCTTGCTATT
 AAGGAACTTGACATCCCAATGTACTCATAGAAAAGTAGAAAATTCAGGGGAAACAAATAGAAG
 AGATATGTCCAAAAATCCATTTAAAATTTCAATTGACAGATATTTATCTCTTTAATCCTGG
 AAGTTGTTTCATGTAAAGGGAAGGTTCAGTAGAATGAAAACAGGACTGTAAAGCTGGGAGGA
 GATTTGGGGTCTCTTGCCCAACTCTGCCACAAAATTGTGTACCTTCAACTTCCCTGCCTG
 TAAAATGAAGTCAGTGGTCCCTAAAATGAAAATTGGACTAATGGCAAATTAGCTAGCATTT
 GATTATACATCGATAGTTTCCCCCATGTGATTTTAACGCACAAAATCAGATTTTGGAAACCAT
 TAGGTAGGTGATCTTATCTTATGAGTTTATCTTTTTCACATGGGAAAATAAATTTGTTT
 TGCTGACCGAAGGCTGTGAACTTAGAACTGGAAAAGAGACATGTAAATTGAATAACCCATC
 TCTATGGGGCAAAAAATGAACACCTTTAGTTTGTTCAAAGATGAATGGAAAAATAAA

CATGTT

SEQ.ID.NO. 473 HG1015065N0 PLT00014333K04.contig
 CCGGCCACTTGAGCTGCTCTTACACTGTATACCTGTGTCTGAGTACTCAGTTCATCCACC
 AGTCAGGGTCTGTGGACAGACTAGGCAGGTGGTGCCCCGTGTGAGGAGCAGCGCAGTGG
 ATCATGATGGAACCCCTCGAAATGAACGTGAAGTGACTGCGCAGTAAGTAATCGGTGCC
 GCTGGGGATTTCCAAAGTTTGAGGGGATTTTCCAGCTAGGGTTTCATCATGGGACAAACAGT
 TATCCCCTCAACAGCAACAGTGTATTGAAACAATTGCCCTAAAGCTAGCAGAGCCTCAGTC
 TCAGAGATCCAAATTAAGGACCTGATGCAAACTGTTGTTTCCATAAACCATGGTTTCCCA
 GAAGAAAGGTACACTAGACCTAGAGCTCTGGGAACAAGTGGGAAGAAATCTTAAACGACGT
 GCGCAAGGCAGTAGTCCCAGTAACATCTTAATGTTATGGTCTCTAATAGGGTTGCTTTG
 GCCCTGCTCTTCATAGAAGAGCCCCAAAAGATAAGGGAAGAGGAACCACTACCCACTTTTA
 ATG CCTCAGCCTCCCTCAGCCTTGCCGTCAGTGGCTAAAGATACCAAAAGAGGAGACGGAG
 GTTTTTCCTCCCCCTCCCCCAATGAACCTGGGAAGAAAGACAAGCGATATACTACAGTTATG
 GGACCCCTGCCCTTAGGCAAGCAGCATTTGGAAGGGAACTCTTGGCCTGCCCAGTGATGCAA
 GATCAACAAGGCAATTGGGTGCATGAGCCCCATTACATTGAACACTTCTAAGGAAATAAGA
 AAAAGCATTAGAGAAAATGGGGCCACTAGCCCCATTCAAAAGAGGATTAATTGAGGCAATA
 GCAGATAACTACCATATGACTCCATGGGACTGGTCAAGTGTAGCTAGCTAAACAACACTTTAGGA
 GCGGAGTCAATACCCCTCTGGAGGGCAGAAATATGATGAAATTGTGTGAACAACAGGCTAA
 CCAGAAATCAATTGCCCTGGACAAAACATCGGCTGCTATGCTCCAGGGGAGGGTCCCTATG
 TTAAATGTACAACAACAATTAAATTTTGTCCCCCAAGCCTATGCACAAGTGTCTTATGCA
 TCTTCAGGGCTTAGGACCAAAATCCCCGAAGGCAGAGTTCAACAGGGATCTTTTGTAAATG

TTCAACAAGGGCCACAGGAGCCATTTATTGAATTTATCCATCAGTTAACCCAGGCAATTA
 AGAGCACACATGGAACATCGACCATTCCACGGGTATCTCGTATAACCCCTCAAGGACAAGC
 CATAGTGAAACGTGCCCATTCACCGCTTAAAAATATGCTTTAAAAAAGGGGGAATATGA
 ATAAAGGACCCTACAACACTACTAGCACAAAGTCTTATTCACCCCTTAATTTCTTAAATTTAG
 ATAATAAATTTCAATCAGCCATAGAAAAGCACTTTGCTAAAAACCTCCCCAAGATACAAAAGC
 CTTCAGTGTATGAAAGATGTAAATAGTAATCTATGTGCAGTCCACATGATTTGCTAA
 CATGGGGAAGAGGATATGCTTGTGTTCCACATCCCCCTCAGGTCCTCTTGGGATTCCAGTGC
 AATGCATCAAAACCATACCATGGCATGGCTGGGACCCCAATGCAGTACTGGAAATGAAGAT
 GTGAGCCTGTAGGACCCCGCAGCCCCGGACAAATGCAGCTTCCTCGGACAACACAGGTCCCCG
 GATGGGGAATGTGAAATAGACAACTGGGGGATGCTAAATAAGACAACTCAGGAGGCTAAG
 TGAATCACAGACACCATTCACCTCCAGATAATCTGTTCCCTTGCTATGCTGTCTGTTGTACA
 TTGCAACTCTTGTAAAGGATGCAAAGCTGGAACACAAAGCTGTAAACCGCTGCACCTGTCAAA
 CCTGTGCTGCACACATCTGTACGCTTCAATCAACTAAACCTGATGCAAAAACACAGAAAA
 GGGAGAGATGTAGGGGATCACTCAGAGTGGTGGGAAAAAACTATAGGGAAAGGACACAAAC
 CTTCTGAAAGGTTGGAAGGTTCTGCAGAGCCCCCTGGGGAGATAGCTGGAAGCAGCTGTT
 CTATAACCCCTGAGACAGAGGGCAAGGAGTAGATACAAGGGAGTGTGGGGGGAATTGATCT
 TAAACAAATCTTGTATGTTGACCAGGAACCTGACCTTTGTATCATCCCTGCATGTGAAGTTC
 CCTGAAA

SEQ.ID.NO. 474 HG1015067N0 PLT00014333K08.contig
 TCATTTTTCAGGCCTGAGTTGATCGTTAATCATCTTAATTATGTTCAATTCAGGCCAAC
 AGGAGAACCAAGACCAAAACTTTATTGTCTCTGCTTTCATTTCTTGTGATGAACCTCTGGA

CTAAGCACACATCTTCCTTGTTTATCTCTCTCAAAGGAGTGTGGAGTGTCTTCATCTGGAC
ATCCACGGGAAGAAAGACATGAGGGAATGCTGTGAAGAGGAGACAGGCCCCAGATTG
GGCAGGAAGTAAACAGTTTTTCAGGCTGAGGCCAATCTGAGCAGGAACATTCCAATATTTT
TTCAGCTACGTTGTCCAGCACTTCACCTGGTTAACCTTTTATGTCCACCATTGTGGATT
TCACAGCTACTTGTCAATGGTGAATATTGATCATCATCATTTACTAGCTGCTACCA
TATCCAGCTACTCCTTGCAATGTTGTTCAATTATTTCTCAACACTCAGCATATTTGCAAT
ATGTTATGTAATATCACAGACAAAGGAAACTGAACGCAGAAATGTTTTATTCTTGCCAAA
CATCATGAGGATGAACAAATGAACCGATTGAAACCAGGATTGTCTGATTCCAAACATC
TCTGGTCCTTTTTCACTCTGATATGCTGCAATTAAAAAGCCATTTCTAAGACTGTACGT
TCCAAAGATTGTTTCTATAGTTGGTTGTGAGAGAAGTTTCTCTGAACTTGCAGAGCACGGG
AAACCAGGAGGAGGCTCAGGGTGCTCTCCACACGTGTGAAACCGGCTGTGGGTGTG
CTTTCCCTGTGACTGCCATTTGCCACTGATGATCTTCTCTCTGGAAGAGTAAGAGGG
GGAAGGACACAGTTTCAGTGCTTCCCATTTATAATTAATTAATTAAGGCAGGTGC
AGGGGCTCACACTTGTAAACCCAGCACTTTAGGAGATTGAGCAAGGAGGATCCCCCTGAGC
CTGGGAGTTTGAGATCAGCCTGGGAAGCAGAGTGAGACCATGACTTAAATAAAATTTT
TTGTGTTTGTCTAATTTTACTTTTGTGACAGGGTCTCGCTCTGTGCCCAGGCTG
GAGTGCAGTAGCGTGATCTCAGCTCACTGCAACCTCTGCTCTGGGTTCAAACAATTCTC
ATGCCTCAGCCCTCCCAAGTAGCTGGGATTACAGGTGCCTGCCACCATGCCCAACTAATTT
TTTTGTACTTTTAGTAGACGGGTTTTCACCATGTTGGTTGGCCAGGCTGCTCGAAC
TCCTGACCTCAAGTGATCCGCCCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGA
GCCACCGCCCGACCAAAAAAATTTTTTTTTTCAATTAGCCAGACATTGTGGCATG
TGCCTGTAGTCGAGCTACCATGGGGCTGAGGTGGAGGATAGCTTGAGCCCAGGAGTT

TGAAACCAGCCTGGCAGCAAAAGTGAGACCCTGCCTCAAAAAATAAAAAAAG
 CTATTTCTTACAGATAATTCATTGAGTGGAAAGCAATTTAATTTTACGTATAATAAGTTAT
 TTCAAAAAGAATTGTAGTTATGTGTGGCAATTTTTCATAAAATACCTGATACAAATCAAGA
 TCTACCATGGTAAGTGTGCTCCAGAGATACCCCTTCTTGACCAAAACAAGCAGGCAGGG
 ATTGGGAGGCCCTTTGGGCTGGGCTCCAGGACACAATGACTAGAAAGATGATGCAAAAGTGC
 ATGACAGAGTAGAGAGCTACACAGTTTACACATATTAGACTCTGCACATATTAGAGTCTA
 CACATATTAGACTCTGCAAAAGTAGTTCTTAAATGTTAGCTCACATGCAACCCAGTCACG
 TTA CTGAGCAGATTCTGGCTCCAGTCTCTGGTGATTCTTG CAGGTGTCATTGACCA
 GATTAGGAGAAATCATCTCAAGGTCTCAAAGCCAAATTTAGGACGCAGAGAGAAAGAAT
 CCTAAGTGGTTGAATGGGAAACTGATTCTCCTCTTGAGCACACCTTAGATTCAAATAATT
 GATTGGAGAGATCATTCATACAAAATCTTCTGACCTGCAGAGTTTCTACAAAAGTC
 TTGAGAAAGAAAAAGAAAAACATATATTTTGGAAAAAGATTGAAATTTAATTATAATTTG
 CAAATTATACTGTGGAAAGGTCAACAAAACAAGCGTATTGCGGATGATCTTATAAAAA
 GAACCTTTTGTAAAGTAAACAATTACCTTCTGATTTGTCTACTCTGAACCCCTTGATGTTT
 ATTTAAATTGGATGTTCTCATAGCAGGACACACCTGATACCTGGCTTGCTAAACACTTA
 GAAAGTCAAAATTGTCTAGTTATCAAAGCCCAGGCTTATTTTCTAAATAAGTATGGATGGT
 CTTTCAGGAGTCATTTACAGTGTGATGTGCAGAGAGAAAAACAATACAAACAACAAACAA
 ACAACAACAAAACAATCATTTGTTCTTAAGTGTGTTTTCCAGCCTATGTCTAACTTTCTA
 GCAGATAATGGC

SEQ.ID.NO. 475 HG1015069N0 PLT00014333L13.contig
 GTGATGGTTGTCGCTGGCAAGATGCAGCAAAACAGCCTTTCTCCAGACCCTGGAGAGTCTG

GAGACTGCCGGACACAGAGATCTGGAGCTGAGACACAGTTTCCTCTATATGCTGGAATAC
CAAAATTCCTTTGCTGGAGTGAGATTGTCTTTGTAAACTAAATTGCCATGTGACGTTTTTTC
TCCAGAAAAACAACAACTTCACTCTGCAGGTCTCTGCCAGAGCAAGCTTTGCAGCTGCAGAT
GCTGCTCACACACCCCGAGGAAGAAAAAGAATGCCCTCCTGCCAGCAATCTGTGAGCAATC
TTGAATCCTGCCACCTTGCTCCCAGCCACTGAGTCCAACTATGAACCTGGAGCATAATTGT
TTAGAAATTCTTGATGCAGTCTATTCCAGCAGACCAGATCTGTGATCATCTGATGGCC
ACTCCGGA CTGGGAGCTATACACTGACTGAAGCAGTTTTCACAGAAAGGAGGACAAACGCCAG
GCCAGGTATGCTCCAGTGACCCCTAGACAAAGTAGTAGAAGCCTATGTCCCCGCTGCTGGT
ACTTCAGCACAAAAAGCTGAATGGATTGCTCTCATCAGGGCCCTTAAACAGTCTCAAGAG
AAATGGGTAAACATTTACACCTACTCCAAGTATGCCCTTCTTGGTAGTGTATGCCCGTGGG
GTCAATTCAGAAAGAAAGGGGATTCTTAACCTCAAGAAATAAAGACATTAAACACCTGCTG
GAAATTTTAGCTTTGTCAGAGGCAGTTGTCTCTGCCAACTCAAGTTGCTATTATGCACCTGC
CAAGGCCACCAAAGGATGACTCCCCAATAACCAAAGGGAATCAAGTGGCTGATAAAGCT
GCAAAGCAGGCCACTCAAGAAACATATTTACTTTGGGACCTTAATACTCCATTTAAACTTG
TCAGAAATTTAAACTCCACTACACTGAATGAAATGAAGAATGCACACATGAGTGGACATTG
ACCAACACAGACCCTTGTTCCTCATGGAAAGCTAACACTCATGGGATATTACTCTCTGAG
GCCCTGGTATACCCAGTCCTGAAAGACCTGCATGAGGGGACACACTATGGCAGAGATGGA
CTGGCTCACCTGGTGTAGCCATATCTCAAAAGGGCCACACCTTCAAAAGATCATTCAAAAG
ATTATACAAGCATGTTATCTATATACCCCAAATGATTCAAGAACTGCATGCAACTCTTCC
CAGAGGAGGATACAAATACAAAGGAATATGACCATTTCGAGGACTGGCAGGTTGATTTTCACC
CAGATGCCAAAAACACTGGGATTTATAAGTACCTACTAGTGTTTGTAGACACATTTTCTG
GAAAGGTTGAAGCTTACCCGAGCAAGAGATTGAAGCCTATCCGAACAAGTCTGAGAAAGC

TGCAGAAAGTAGCTAAGGCTCTGCTAAAAAGAAATCAATTCCTCGATTTGGACTTCCTTACTC
 CATAACCAAGCAACAATAATCCTTCTTTAATTCTGAGATCACACAAGGATAAGCCAAGC
 ATTACAGATCAAAATGGAAATGACACTCTTGTGGAGGCCACAATCAACTGGAAAGACTGA
 AAAGATGATCCACATGCTAAAAACAACAACAACAACAACAACAACAACAACAACAACA
 CAAAAAACCAAAAAACCCCTGGCAAAACTCTATCGGGAACAGATTTAAAAATGGTAA
 AGGGTTTTACCTATTGCCCTGCTATGGATTGGGGTAGCACCTGAAGTGGGCTTAAATTA
 AGCCTTTTGTAAATTGTGTGGGAGACCCCTTCTGAAGTTCTTGGCCAGGGACCCCACCC
 TTGGACTTAATAAATTAATCAAGAAATTAACAAT

SEQ.ID.NO. 476 HG1015070N0 PLT00014333M01.contig

GACTGGCCGCCGTGCATCCAGCCAAGCTCCTGGCGACGGATATCGGAGGCCAGGCAT
 TCGAGGGACGTGCAGTTGCATCACGGAGAGCCCTTGGGAACCGTCTGGGGTCTTGGCGGA
 GACAATTTGTGTTTCGGGGCAGCGCCCTTCTACATCCAGAGGCCCGCTTGCCCGGGT
 TACCGAAGGCCCGCCACCGACGCCCCCGCCCTGACCGCCGATGCCCTCCCCACGCCGGC
 TCGGTCGCCCTCGCGCTGCCGTGAGAGATTGCGGGATTCAAGGTGACTTGATTTTC
 GCCTGATCCTGTGCGCAGGGTGGGGTTTCTTGGCATTCCTCAGGCCGGGGATGCCGACG
 TCTGGGTGTGAGGTGATGAGCTAGGGTTCTGTGGCACTCCGGCTCTGCTCAGGGCGAC
 CTGCGTAGGGGCCTCTGGGGCAGGCCCGCTTCTGTCAAAATGGCAGGCCCCAGCGTGCG
 CGGGTATCTCTGAAGTTGGGGTTTAAGATTCTTGCCCCGTGAGCGAACGAGCCTGCCGT
 ACTTCTGATTTTGAACACCTGGTCTGGGCTTCCCGGGCGAGCGGTGGGTGAGTGCCGG
 CCTCATTTGCAGCCTGGGGTACGACGTGGGCGGTTGTGTGTAGGGGGTTTACATTGGCC
 GTGGATTTTTCCTCTTCCGGATGTGTTTGTGCTTTTGTGATGAAGAGTTCCTCCGTGGTTT

TGTAACCTAATAAATGCATAAAAATGCAGCTTGGGTTGTAGACTTTGGAGAAATAATTAGAG
 ACCTTTAGGGCTTTAGGCTTTCTCCAAGAACGGAGCAAGTTTTCAACCATGGCTTTCTT
 CAGAAAGACTGCAGCTTTTTTGAAATCATCTTATTAAAGTGGTGGATAAAGTTGCTTGTCT
 GATCTTTGAGCAITGGCACTGGACGCAAGGAAGCTTCAGCCAGTCCTAGTTACTGTGTTTGC
 AAAGCACTTTCTTGTTTTTTGGAATCTTGTCTTATGTAATCAACCTTAGGAAGAACCCACA
 CCTTTCTGCACTCAAGCACAGCGAGAGTACGTTATGCGTGAAAACAAACAACTTTGCA
 TTAAAATCCATCTTTCTCCGAACCACTTGGTCACGGTTGTGTTCTCATCTGCTTTAGTGT
 ATTTGGTTTTGAAACGATATACCTGCTGTATGTCTAAACAACTTTATAATTTTCAGGAACT
 TTTCTGTTTCCCTTTCTTTTTTTTGAGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTG
 CAGTGGGGGATCTCGGGCTCACTGCAACCTATGCCTCCCAGGTTCAAGCCATTTCTCCTG
 CCTCAGCCTCCTGAGTGGCTGGGATTACAGCGCGGCTACCTCGCCCAAGCTAATTTTTG
 TATTTTTAGTAGAGACAGGGTTTCAACCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCT
 CAGGCAATCCGCCCGCTCGGCCCTCCCAAATGCTGGGATTACAGGTGTGAGTCATCAGG
 CCCAGCCTCTATTTTCAGGAACTTTAAAGGCAGAAAAACACTTCTCTGACCTTACCATT
 CTTTAGTGTACTTGCAACCTGGTGTCTTGCATCCAGATAATCTTACTAATCCTGATGCA
 ATGACATTTTTTAAAGGCACTGAAGAGAAAAATACTTTTATTAAAGATATCTTTAAGTCTC
 CTCCCTGATTTCTTGAAAAATCTCATGTCTGTTATTGTTTATTATTATTTTAAACAGGA
 TAGCAAAAACCTTCTATTACTTCTGTTGTCAGGTCAATTGAGCAAAATCTAGCCTCTGCTT
 TATTGAAAAGGACTTTTTTGTAGTTTTCCCTTTTGCTATTTCTTAAAGGATCCAAGACCAAGG
 CAATTGCACATTGTTATAGAGTAAGATTTCAGAGCTGATTTCTCTTGTGATACACTATTGTGT
 GCCAGTGATATGAAATGATATGCGGCACCTCAAAGCTACAACATTAAATGCAAGTTACTGA
 TAGAGAAATGCTATCAGAGTGAAAAGGAAAACTTGAAGATTATTTTGGCTTTTTTAAAAA

TAGCAAAGATTACATTATTTATTATATTTTGCTTTTGAGAGTAAAGTTGGAATAATTAC
 TATATACAAATTGTTTCTACCTAATTATCAAGGAATTTAATCATTAAACGTACATAAATTAAT
 GTATCTAAATTACTCAACTTTGTTGTAAT

 SEQ.ID.NO. 477 HG1015072N0 PLT00014333M02.contig
 GCCGGCGGATGGAGACAGGACCGGCTGCTGGGTTTGGAGGGCAGAGATTGAGGAG
 CCAGTTGCTGTGCCCTGTTCTGGCAAAAGAACAAAGATGCGTTGCAAGCCAGAAATGATGAA
 TTCCACTTTTACGTGCGCTCCGTGCGCATTTCTGCCTAAATGATGGCATGGAAATGGTGC
 TGAGGGATGTGGGCTGAATCGAGAACACAGGGACTTTTGTCTTTCAGGGTGAGAAATGTTC
 CGGTCCTTAAGGAAACATCGTGCAGCAGTTGCAATGGGTCTTCCAGGCGTTCTCCCAAAA
 CATTATTTTTTTCTTTCACATAATTGAGTTTAAAACTTAGCGTTCTGTGCTAGCATTGC
 ATCTTAGTTACACATCCTCTCCCAACCATAACGTCAAATTTAGGAGACGATACATCCCCA
 AGCTTTCCTGCAATATTTTTTGGCAGTAACCTGGTTATTCAAGTTTACTTTCCTTAAACTATG
 TCCTTCTCAAACCTAGTCTCACCTGTTTCGTTACTGTCCGACATTTCTCTGCGAATAACACGT
 TTCTGTAGTTCGTGTTAGAGCCTGGCATAACCTCTCCCGTTCCCTCTTCCCAACAAGT
 CATCATGCAGAGGGGATTTTGGTCTGAGTTAAAAGCCAGGTAATTCCTTTCACCATCTTTTC
 TGTCGATAGTCAAACTACTACAAAGACTACTCAAATTTAGTAGTTTCTTATTTCCCTATA
 GCTTCGTTATATTAGAAATGGATAACTTATATTGCAATACCTTGTCTGTAATGAACACACA
 CCAAGTTTTCATTTAAATTAATCAAGGACAAATGTCTCTGAATTTAAATGTGTATGT
 AGTAAATATAACTCAGCATTTCTTGATTACATATCATTAATAATGAATGGTACTATCT
 GAGCCACCAGAATCAGCAGTGAAAGGTATTATATTAATAAAAAAGGACTAAGAATCAGGACT
 GTTTTAAAGGAACTGGAAACCGAAGAGGATCTACTATTTTCATCCTAGGCTAGTTAACATGA

AATAGTGAAATTAAAAATTGAAAGAGCTTACATGGATGCTTCATATCTGAGCTACTTTGGCA
 GGGATGTTTCAAGAGGATGTAGGTAAGTATGGTGTTCACCTAGCTTTCCAGAAGTCTCAT
 ATTTTGTGTAAAAATTGGTATGCTTATTTCATGGTACTTTGGGTTAGTAGAAGCTATACTT
 CTTTTCCAAGGAAAGACAGACTGTCTTAAGATTCTAACTGTGCATCTTGGTGATAAGAAG
 TTCTTAGGGTTGTTAAATAGAGCCAGGAAGCGTGAAAGCTTTCTCTCTAGATCATGACAAAC
 TCATCTAGTACCTTGCAATTACAAAAGGAGCTCTAAGTCCATTTACTTTGCAAGGAATGAAG
 TTCCAGTGATTAATGTTTTTTGAGAAATTTGTAATTGGGATTGTGGGTCCAGAAGGTTGCC
 TTAAGTAACTGTTTTTCATGACTCCTAGGTCCCTAGATGTTGAGAAGCTCTGGATTGAAGC
 TGTACAAAAAATAGTGAGTGGGCTCCGTAGTCACTCCATTAGCATAATACAGAGTAGGA
 CAGCTTTGTGACATTTCTCAGGCTATTTCCCTGACCTCTCATCAATGCTGACGTTTTTG
 ATAAGAAGGGGTATAGAGACTGCAGCCCTTGGGGTGTACAGTTCAATTAATTACAGGC
 CTTATATTGACAGGCTCAGATTTCATGCTCAGTACCATAAAGGTTGATTTTAAACCACC
 AAGAGGCAAAAGAAACATGACAGGTCGGAGCATATTTTCTCACCTATTCTGGCTTAAATG
 ATGGTACTTACCTGATTTGTGGTTGTAGGTGAAATGTTGGGTGAAAGATTAGCTGCTATA
 AAAAGGATGTCAGGACGTGTAGAGAAATTTAGAGGACTTGGGTGGGAGATGGATGTCTATT
 TTCTAACCCCTTAGTCATGTTTCATTCCGATGCCAGCATAAACTCTATTTTGGAGATGATTA
 TAGAAATAACTTATATCTTCAATTTGTTGTGGGAAATAAAA

SEQ.ID.NO. 478 HG1015074N0 PLT00014333M07.contig
 CGTCCCCGCCGCCGCCCTTCTGTGCCGCCGGGCTGCGGTCCAGCCCCCGCCTCG
 CTATCCTAGTTCTGTGTTGGGTCTGTCTTCTGTCTTGTCCGCCCTCTCCCGCTTCT
 TCCGGGGGGACATGCCGCAGAACTAGACTGCTCAGTGTCCGACGTCTGCTCTGAC

AACCCCGCCGACGCCGCAATTCGCTGCCCTCAGCAAAGGGGTGCTCAAAAACGGAAAGTGAT
 CATTTTACCAAATAGGCATTTTCCAGCTTAAAAAGATGTGCACATTATTTTCTTGTT
 TACTCTTCCTATAATCGATACCGTTTAGTGAGAAAAGCCTTCTCCTTGCTCCATTGTTT
 TGAAAGAATTAATATAGACTGCCCTCTCATCAAAATTTTCCACGATTTTAAATTAAGTTAA
 TCTCAGACTGTTAACTTTTCTTTTCAGCTGGACACATTTAGAGACTCAGTGTATCTTTCT
 CTTCACTCCGGAAAGAGACAGCCACTAAGAGATTAGCTGATTTGAGTATCCCTCTGTGGT
 CGATTATACCATTGCAAATCCTGCTCTGCTTTAGGTTGGCAGTTTATTATTTAATCCAGT
 AATTATTCAAAATTCACCTTGGACAGATTTGTAAAGTAGCAGCAGATATTACTTAAAAATC
 TTTTCCTTTCCTGCTTTTACAAACATGTAGTTCCTTCCTGAAAGGTAATGTTAGTTTGT
 GGGGGTTTGTTTTTTTGGTACTGCTTCTCTACTGCTACTAGGTGGGAAGCTAGTTC
 TTAATTTTAGGGCCTTGAAGTGACACATACACAAAATTATAATCTGCTAATGAAGAAATA
 ATTTATTCAGGTCAATAATAAAGTGCAAGTTAATCACTGTTCCAGTACTATACGATGTA
 TTTTTCATGAGTGATGAACAGGCCCTGCAATGTTTTCTTTACAAGAAATCACAGCAACA
 AGGTTATGTTCTATTGCCCTGTTATCTGAGCTCTGTTTTTCAGACATAATTCATATTCAG
 GAATAAACTAAATCTGAGTTTTATGACTTGTATGGATTTAAGAAATGTAGGGAGGAGTTT
 CATTCTAATCAAAATTCCTCTGACACAATGGTTCACCCTGATTTACCAGTTTTTGACTAC
 ATTGCTGAGCAGAGAGAAAATTGCAGATTGACTTGCAGTATTTTGACTGATTCATGTCC
 TAGTGCTTCGGGGAAATTGCTTAAGGAAGTTGGCGTTGCTCAAAAGTATAGTTGAGCAAA
 ATATAATCCTTCCATAATTCGTGCTTGGGAAAGGCCACTGTATTCTCAGTCTGAAAGAAG
 GGAAATGCTAGTTAGTTGGGACCAACCCTGACAGCCACCAGAAAGTGCTTATTTACTGCC
 AGTACAATAGTTTGACTTAAAAAGCTGCAAGCGTCTTGGGTACCTTGCTCTGCTCTAGCCA
 CAAATATTTTCCCACTTTTCTTCAATCTATTGTGTCTCCTCATTTAGTGAACATTTGCCAGAG

CTGTCAAGAGAAATAACCTTAGAATTTGTATGGTAGTTTACAGTTTACAAATAATATATA
 ATAGACTTAATTTTGAGTACCAACATCCTGTGTGACGAGCTGAAAAAATAGCAGGTTTTG
 TTTTTAACTTTAAGAGGTTTGGAAATTGAACCACAGAGTTGGGAAGTCAGCAGGTTTCA
 AGGCTAAGACAAAGTGTCTGTAAAGAAACAACCATTTGAAAAACAATGTATGCAGATAAGAG
 GGTTTATTATAAGGATATTGGAGGATCTTATAGATCTTGACAACTTGGTGTCAACCTCTG
 CTTTTCTCTCTTTTCTCCAAATAAGATCTCTCTGTGTGTCAGTCAAAATGGCTGAAT
 ATGGCTACCCTACACTTTCCAAATTGACAGGGTTGTAGCTGCATGCAGAGATTACATAGC
 TCTTCTAAGTCCAAAGTTTAAAGTTCCTAAGAGAAAAGCACGTGAATGACCCCTGCTTGGCTT
 GTGTCAAATCATCTCTTGGCCAAGGTGGCTGTTTACCCAATATACACTGCCCATGAGTGA
 AAGGCAATTATCACAGGAGGGGAGATGGATCTTGGGTGTGGGAAACAGGTGTAGAATTCC
 TACTTTTTACAGTTTGGTAGAGGTAATGTTTAGAACCTAAGATGGTGTCACAAAACACA
 AATGCTTATGGGGTCAGTTAAATAACATAAATGAATGAAGCTAATAATGGAGACTGTGG
 CAATGAGAGGGAGCACAGGATCCATCTAAAATGAGCTACTACTAGGCCAAGTGGGATGGC
 TGACACCTGTAGTCCCAGCTCTCAGGAGGCTGATGTGGGAGGATCACTTGAGTCCAGGAG
 GTAGAGGCTGCAATGAGCCATGATCACCCCACTACACTCCAGCTGGGCAACAGACCCCTG
 TCTCAAAGGGGAGAAAAAAAACCAAGGAGCTACTACTCAATTCAAT
 TCTGACTGACTGATGCTGTGGGAATTCAAGCGCAAGGTGACCAATCTTATGTTTTGCA
 AGTGATGTCGGAAACTGGGATTTTATGTAAAATCTCTATTTCATTATGTAGGAAAAACAA
 AACAAAAAGAGAACTATAACTGCAAGTCAGGTTTGACCTATGAAAATGGTAGTTTAGGACT
 TCTGACCTAGATCATCTGTGTGGTCAATTGCCAAAATAATGACATTTTAAAAATCACAGCGT
 CCATAAAGTCTTTTACCATAGAAAGAGACTATGATGCGTGCCACATTAAATGTGAGCAGTA
 TACATGTTTGTGTATAGTGCAATATACACAGGAAACTAAAAAATCATTCATCCATCCATC

TTGTTTCTCTGAAATTGAAACCTTTTGTAGATTCAATTCTTCTTTACTTTGGTGAGG
 TAATTATAAGAGCTGCACTATTTGCGTTAGATTTTCCTAGAGAAACCCCTGCTGTAA
 AATTTTAAAAACACATAGAAGTTGTATGACCATGAACAACATTATTTAACCTAAGTCTC
 CATTTTCTCTCTTGTATATGTCATACCCATGTCTCATGGTTGTGCATGAAAAATCAAGTGA
 AATATGACTTTATGAAGCTCTAACATGTATGAATGCTCAATGTATGTTAGGCTTCATATT
 TTGTTAGAAATTAAGTCTAACACAATACATTACAATACAGTATGATTCAA

SEQ.ID.NO. 479 HG1015076N0 PLT00014333MI5.contig
 GCGCGTGCGGCGTCGGAGCAAGCGGCCCCCTTCTTCGGCCCCGCCAGCCCCGCCGCGC
 CGCCCCCTCCCTGGGGAACCGAGGGAGGAATGGCGGGGTCCCTCAGGTGCCCGCGGACGT
 CATCCAGCCCCCAAGGCTCCAGGACGCAAGGGTCCCCCTGATCTGGCAGTCCCCTGGACCT
 ACCCTGGAGAGCGTGGGGGACAAAGGCCGGACGCGCCCCGCTGCCGCCCTCCAGGAAG
 CCTCCCTTGCCAGGACAGTGGAGTTGCCCGGGTGTGGACAGCAGGAAGGCCGCGCGGGA
 AGGGGGGTGGAAAGAGCCCACTCCTGGTTAATAACTCTGCGGGGTTAACCTAGGGCCGGC
 CATACTCAAGTACGGACTTGAGTAAATTTCAGCAGGAAAGGAGCAGGTTGGCGCCGCT
 AGAGGGAGGCTTCAGGACTGAGGGTCAGAGGCCCAAGGCCCTGGCAGGGCCCTTCCCTGGT
 CAGCTCCAGGCTTGGCCTGATGCCCCGTGGGCCAEGGGCCCTCTTCCAGCCTCAGGCTGAC
 CCCAGAGAAAGAGGACAGCCACAAAGCCCCGGAGGCCGGCTTGCTGCCCTGGTCAATTGCTG
 CAGAGCCTGGAGGGGCTGCGCTCAGTCCCCGCTGACCGTCCCCCTCCCTTGACCCCTGCC
 ATAGCCATCTGTGAACCCAGGGGAATTATTCCAAAGGGACTTCCACTTCTGTGATAATCT
 ATGATTCTATTTTCTAAAAAACAACCTCTAAAATGCTTAATGTCAAACATCTGTGATCA
 GCTGAACAGGAAAAGGAACAGTAAGTACAAAGCAAGGCAACATCATGAAAGGGAGCCCTG

AGGAATTTCTCCCTCCTTCCCTGAGCCGGGTGAGGCCGTGACTGCAGCCTGACTCACAGG
TTTTATTTCACATTATTAACTAAAGATGGTGAGGATCTGCTGGTTCTCCCCCAGGCCTGG
AGGCTGCCCTACCATCGGTGGCTCTCATCGGTGAGCTGAGGCCACAGTTGTTATAGTCTCT
GACTTCAAAGAAAAGACCGATTTCATGAACAAATTGCCCTGTCTCTCTCGACGGTGGCCTCG
GTTTTTGTGTGTTTTTCTGTGTGTTTTTGAGACAGAGTTTTTGCTTCCAGCCTGGGCG
ACAAAGTGAGACTCCATCTCAAAAAAACAAAAACAAAAACAAAAACAAAGAGTTACTA
TTCAGAAATATATTAAGAATTCCTACAACTCAACAAAAAAATAATCCAACCCAATTAAA
AAATGGGCAACAGACCCCAATAGACATTTCTTAGAAGACACACAAATGGCCAATAAGCTC
ATGAAAAACATGTTTAACATCACTAATCATTAGGGAAATGCAAAATCAAGCTAAAATGAGA
TACCACGTCACACTCATTAGGTGTTATTTTTAAAAATAACACAACAAAAAAATAAGA
AATGTGGCGGAGGATGTGAAGAAATTGAAACCTTGTGGACTGTTGTGGGGATGTAAAA
TGGCGCAGCTGCTGGGGAACACAGTATGGGAGTTCCTCACAAAACTAAACAGAAATTGCCA
AACGACCCAGCGATCCTGCCCTCTGACATATACTGAAAGAACTGAAATAAGGCGCTGGAA
CAGACACTTGTAAACCCATGTTTCACAGCTGAGCCAAAAGGTGGGAGTAACCCAAAATGTCCA
TCAGTGGAATGAATACTGGCAATAAAAGCTGTGTCTGTGAGGCCACTCCGTTTCACACAGCT
CTCCACACGGGTGAGGGCACAGCCTTCTCGTCTCTTTTCCCTATGCTGCATCTTTCAGAGCTA
GCTATGATATGATGATGATATGTACATATCGAATATAC

SEQ.ID.NO. 480 HG1015078N0 PLT00014333N05.contig

TATGGATGCTCTGCAAAATTCTCTCTTTTTTTTGACACTTTTATCTAGACCGGAAGAGTTC
AATCTCAAATACTCTCCCTGCACCCCTGGTTTTTCCATCACCCCAAACTCCATTTCTTTC
CTCCCAGTCACACCAAATTTCTGAAGTAGCCCAACCCGTGCTAGCCGTCCCTTGCTCCTAC

A C A C T C C T T A G T G C C C C C T C A T G G G G C T T C T G C C C C T A A G T C C C T G G A C T G A C C A C C C A G
 T C C G A A G G G G T T T C T G G T C C T C A C C A A T C T G T G G C T C C T G G C T C T T T C T G A G A C A C T G A
 G A T C C C C C T G C C C C A C C T C A T T C T T G C T G C T T C C T A G A C G C A G G T G T C C T C A G G C T T C
 T C C A G G C T C T A A C C T T C T C T T G G C A T T G G C T T T T C C A G G C T G T A C A T T C T G T A C T G G T G C
 C T T C T G C A T C C T C A T C A T G C T C G G T C T C T C T T C A G C C T C A G G C C C A C C T G T C T T C T
 C G A G A C A T T G C A T T T T G C G T G T C C C A C A A T A T C T C A A A C A T G T C C C C A A A C G A A C T C
 A T C C A T T C T G A A G C T G C C T G G C T T A A C G A A T G G T G C C A T T G T C T C C C C A G A A G C C T A G A A
 A G A A A T G T G G G A G C C T C A G T C C A T G A C C A T G C C C A A T G C T C C C C A C C G A C C C T C G T C T G A
 G C T C T G G A T G T T T C T G T C C G A G G T T G C C A C C C C A G C A G C C T A G A A A G A A A C A T G G G A T C C
 T C T C T C C A T G A C C A T G C C C A A G G C T C C C T G C T G T C C C T G G C C C A A G C T C T G G A T
 G T C T C T G T C T G A G A C T G C T T A C T C C C A G C T G G T C T C C C T G T G T T C A C C A T G T C C A C C T T T
 A A C C C A T T G T C C A C A C C G A A G A T C C A A C C G T A A A C C T G A T C C A A T C C T G T T T A A A C T C T T
 G C G T G T T T C C C A T T G T C C T T G C T G G A A T C C C G A G T C C T A A A G G C A G G G C T C A G G C C C C T C
 T T G G T C T G G T T C C T G T T T G T G T T T C T A T C C T T C T G G T C C T T C C T C T C C C A G A C C T C A
 G T T G C C T C C T C C A C A C T C T T C A T A T T C T A A T C G T A C A A A T A C T T T G T T T T T G G C T T T C T G
 A T T C C T C A A G G T G T T C T G T G T G T T T T T C C T T C T T C C T C C T C A C T T T C T C T G G C T A T T T T C
 T G C C C A G C C C T T T G G C T C T T G G C T G T A C A G C T C T C C C T C A G T A A C C T T C T T T C C C A T C C
 T A G C C A A G G C A G G G C T C C A G G T T C A C A C T T C T G T A G C A C C C T C A C T T T C T A T A G C A T A G A
 C C C C G T G A C A C T G T G T T G A A G C T G A T T T T A C A T T A T A T A G C C C T G C T G G T C T A T A G C T
 T T G A G A A G A T T T G G A T C C T C A G G G C T T G G T C A G A G T A T G A G C C A T A G T A A A C T C T C A A T G
 A A T G C T A G T T G A T G A G C A A A T A A G T G T A T T C A G A G G T A C T G G A G T G C C A A G C T T T G T A T
 A T A T G C A C G C A C T C A G A T A T A C A T A T G G A T G C T C T G G A C T C T A C T T A C A C A T A T A T G T A

TATGTGTATATACATGTATATAAATATATATATGTTAAATATGTATATATGGTCCATGTATA
TGTGTGTACATATATATGTATATACATATTTAAACCACATTATTGAGGTATGATTGATAT
GTCAAAAGCTGCACATATTTAATGTATACAACCTTGATGAGTTTGGGGCTATGTATACACC
TGTGAAACCATCACTGCCATCAAGGTCATAAACATTCAATGCTTTTAAGAGCAGGGCTTC
CCAGAAATTAGGAATTTTCATAGACCATTTAAATTTGCAGGAAAAATGTAAACTACATTTCAGG
TCGCCATCTTTACATTTTCTAAGTGAAGATGTTGATTAAAAATAATGTTTGTTTTTCCTC
ACCACAAATCCCACAAAAGACATAGCATTTTAAACCAAAATGGGGCTAGGTGTGGTGGCTC
ACGCTGTAAATCCAGCACTTTGGGAGGCCAAAGGTGGTGATCACGAGGTCAGGGAGAT
CGAGACCATCCTGGCTAACATGGTGAAACCCCGTCTCTACTAAAAATACAAAAAAATTT
AGCTGAGCATGATGGTGGTGCCCTGTAGTTTCAGCTACTCAGGAGGCTGAGGCAGAAAGAA
TGGCATGAACCCAGTACGCAGAGCTGGCAGTGAGCCAAGATAGCACCACTGCACCTGCAGC
CTGGGCAACAGAGTGAGACTCCCATCAAAACAAACAAACAGCAAAATGCCTACTGAAAA
TCTCCTTGCAGTACCGGCCTACAAATGTACCGTTAGGAAGCTGTTCTCTTCAAGAGCAGG
CTTGGAAGCTAGAGAGGACTGAGATTGCAGCTGTGGTCTGAAGGGCTGTGTGTGAACGGG
GGAGCCAGAGAGACTTGGATCTATCAGTCAGGGTTTCTAGAGAAGCAGAACCAATAAGA
AGGAAAGAACAAAGAGTTATGATGAGGAATTGGCTCACATGATTATGGAGGCTGATGAGTC
CCAAAGATCCGTGGGATGAGTTGGGGAGCTGGGGGCCCCGTGAGAGCCAAATGGTGCAGTGCA
GCAAGGTGCTACGGACCCCGGCAGACAGGTGTGGATGATATCTTTTGCCAACTGCCCCTTG
GACAACTCAGATTTCAACCACGGTCAACAAATAGAGTTACCTCTGACTTCAGTTTGGAAA
TTTTTTTTTTTTTGCAGGGAATTCTTTGTAGATTTTCAAACTGCTGTTGATTAGTTTGTGTC

A

SEQ.ID.NO. 481 HG1015080N0 PLT00014333N11.contig

AGCCCTCAAGGAACTGAGAAATTTCTCCCCACTTTGTTCTGAGGGGTCTCAGCTACTC
TGGTATTTAAATAAATGGGTTTGGAAAAATAGTTACTGCCCTTTAGTTGATGACTAAA
ACAGAAGCCAAGAAGTGTGCAAAATTGCAAACTGACATGCATGAGCCAAACATATTCCTG
AAATGACAAATGTTCAAGACGCAGCTAAAGTCACCACTCTGGCGGTAAGCGTGTTACAGAG
AACTAGATTTCTTTCCGGCATCCGCAACTTGGCTGGAGTGACCAAGGAGGAGTTGAAGAG
CGCTGGAGATGCGAGGGTCTCGTCCATCGCCCCAGGCTGGAGTACAGTGGCATGATCACT
GCTCACTGCAGCCTCACCCCTCCAGGACTCAAGCAATCCTCCCCACTCAGCCTCTCAGGAG
CTAAGGACTGGCTCAATTGACTATAAAGAAATCGAGAAATGTCAGCTGACCAAGGCAACCAGG
AGACGCTTTCCTGACTTCCACTATGCACGTGGGTGCAATAATTGTGCTGTGAAGTAATG
AAGAACGTGCTTGCTGTAAACATCCAAACGCGTGCCACCATTTCACAGATAGTGTCTT
TGGGAAAAGGTGTGGGTATAGATGGGGAATGGTCAGTCTCTATGAATATGGGGCTATAAGAC
AGCAAGGCTAGAAAGTACTGTGCTTTCAATTTTAAATTTATCTATTTTTTTTTTTTT
TTGCACTAATGGTTTGCAATTCACATTGAGGAACTGAGACTGTTTTCAGACTTTTGGCCTA
CTGATGACCTTATGTTTATAAAGGATAGAAGTATAATAAGTTGTTTAATATAAAACATA
CTGAAATTTTATTTTCTGTCTCCTTGGATATAAAGTTTCCACTTATCTTGGGTTTAAGTA
CCTACTCCTCCACAGAGAGGGACCTTGTGCATGTAGATTGATATTTCTCAGGATGAAAGCA
GTACCTACTACAACTAGGAACCTACAGAACAGGAAGACGATGCTTGTTCCTGATGGCATG
TAAATCTAGATTCACTTCTAGCTGGAAGGAGGCAGGCCAGAAATTCAGAGTTGTAGGTCC
TTGGCCTGGGTGATGTGCTACAACATGAACCATAAACCTCCATCACTTATCTTGGCGGA
TAGAAAAATTGGAACCTTTGGAGGAAAGTTTCGCATCAAGCACTCAGACCCCAATAAATACTC
CCACTCCTTGACAGAGCTACCGAACATTTCTGCACTCCATTATCATAGGCTACCCCAAGGATGC

TTGGAATGTGCTGGCCGGAGCCACCATCAGCACTTGCTGATGACTCTTTGTGGTACCGCC
 ACTCACTTTCTGTACTGTGTTCTTTCTTCTTCTTCTTCTTTTGTGTTCTTCTGCCCAGCGAC
 ACAC TTGTGCAGACTGACTTGGTGCC TTGGTGCC TGT TATGGACTGAATGCTCGTGTCCCT
 CCCACATTTCATATGTCAAAAGCACAAATCCCCAATGTGATGGTATTTGGAGATGGGGCCTCT
 GGGGGGTGATCAGGTTTAGCTGTTGTAAAGGTGAGCCCTCCGCGTCTGGATTAGTGCCCT
 TATAAAAAGAGACACAAGCAGCTTTTCTCATTTCTCTTCCACCATGTTAAGATACAATGA
 GAAGACGGCCATCTGGAAAACAGGAAGCAGGCCCTTCACTGTGTAGGCAATCATAACAC
 AATGAGAAAGTATTGTGTGGCACTATCTGTGTCCAGGAGCAGCTGGACCTCCCAGAAAGAC
 ATGGGGTCAGCCTCCACCTGGATCCTGAACTTCCCAGCCTCCATAAGCATGAGAAATAAA

TG

SEQ.ID.NO. 482 HG1015082N0 PLT00014333O03.contig

GCTTCTCAATGTTTGTCACTTTTGAGGTGCCACGGTGTCTTTCTTTGGCCACCATTTCC
 TGGTTTCTTCC TTGTCTGCAGATCTTACTAGAGTTAGGCCAAAATCAATTTCAGGTTCAGC
 TCTAGTGGACTGGGAACAATCTCCAGCAGGGAACCTCCCTGGACAGTCTGGGGACAGCTGG
 GACGATTCCTGAAGCAACTTAGATAGGAAAAATCACGATGAGCTTTGATTTAGGGAAGAA
 CATTGTTGGCACCCCTGAGTCCACCTGCCACCAGCCGCGGAGGCGCATGTTATAAACACAC
 CAGACACGCCAAGGATGGGAACATTGGACCTAATTTAGCAACAACAATCCTTTCCCTCA
 TGCATACTTCTTAATTACAAATGAGATTGTGGTAGGCTTAGCATGGGTAATAAAATTG
 ATTCTGTCCACAAAAATTGGAGATGATAAATGTGTTCAAGAAGGGCTTCAAAATGCCTAGC
 TGCTGATTAGAAAAAGGCAAAATAATTGTGAGATGGTGACTAATAGATGGAAGGAATTGTG
 GTAGAAATTGGTCTCCTTGCAGAAATATCACTGGTCTGATCTTTTTC AAGATAATTACCATC

AAATGACTGTGTGTTTCATGCTCTACCCCTTCAGTTGGCCGAGGCATTCAGGGAATCCCCA
 CATTGACTCGGCTGTAGGTGGCTCTGGCTTCTCTCTGTTTGAAGCCTGTAAACGGAGCCC
 CTCTGTTTCTTTTTTCTTTTTTAAATTAACITTAAGTICTAGGGTACATGTGCACAA
 CGTGCAGGTTTGTACATAGGTATATAATGTGTCATGTTGGTGTGCTGCACCCCATTAACCTC
 GTTATTTACAGTAGGTATTTCTCCCTAATGCTATCCCTCCCCCAGGCCCCCAACCCGCAACA
 GGCCCCAGTGTGATGTTCCCCGCCCTGTGTCCAGGTGTTCTCTTTGTTCAATTCCAC
 CTATGAGTGAGAACATGAAGTATTTGGTTTTCTGTCTCTGTGATAGTTTGTGAGAAATGA
 TGGTTTCCAGCTTTATCCATGTCCCTGCAAGGACATGAACCTCATCCCTTTTATGGCTG
 CATAGTATTCGGTGGTGTATATGTACCAATTTTCTTAATCCAGTCTATCATTTGGTGAACA
 TTTTGGTTGGTTCCAAAGTCTTTTGCTATTGTGAACAGTGCCACAATAAACATACATGTGCA
 TGTGTATTTAATACTAGCATGATTTTATAATCCTTTGGGTATATAGCACCCAAATAATGGGAT
 CACTGGGTCAAATGGTATTTCTAGTCCTAGATCCTTGAGGAAATTGCCCACACTGTCTTCCA
 CAATGGTTGAACTAATTTACACTCCCAACCAACAGTGTAAAAGCGTTCCTATTTCTCCACA
 TCCTCTCCAGCATCTGTGTTTCCCTGACTTTTTAATGATCGCCCACTCTAACTGGGCATAAG
 ATGGTATCTCATTTGTGGTTTTGATTTGCAATTTCTCTGATGACCAGTGATGATGAGCATTT
 TTTCATGTGTCTATTGGCTGCATAAAAATGTCCTATTTTGAGAAAGTGTCTGTTCATATCC
 TTTGCCCAATTTTGTATGGGTTGTTTGATTTTTTCTTGTAATTTGTTTAAAGTTCTTTG
 TAGATTCGTGCATATTAGCCCCTTTGTGCAGATGGGTAGATTGCAAAAATTTTCTCCCCATTCT
 GTAGGTTACCTGTTTACTATGATGGTAGTTTCTTTTGTCTGTGCAGAAAGCTCTTTAGTTTA
 ATTAGATCCCAATTGTCTATTTTCGCTTTTGTGGCATTGCTTTAGGGTGTTTTAGTCAATG
 AAGTCCTTGCCCCACACCTATGTCCTGAATGGTATTGCCCTAAGTTTCTTCTAGGGTGT
 ATGGATTTAGTCTAACATTTGTGGATGAATGAGTCTTCGTGTAGCAATTTTAAGGAGGGA

GATGTACAGACTCATTACAGATGAACCTGTTTGCTAAGCACAGTACTTAAACTGCCATTT
TGATAGATTAACTAGTAGTTAGAGGACATCTTCAAATTGTCAGGAACAAGGTTTGAA
TAGAAGATTTGTAAGTCACTGTTCTCAGTGGCTCATGCCTATAACCTCAGCACTTTTGGGA
GGCTGAGGCAGTCAGATTGCTTTAGGCCAGGAGTTTGAGACCAGGCTGGCCAAACATGGTG
AGACCCCATATCTACAAAAAATTGAAAAATTAGCTGGGTGTAGTGGTGTGCCTGTGGT
CCCATCTGCTTGGAGGCTGAGGTGGGAGAAATAGCTTGAGCTCAGGAGGTTGAGGGCTACA
GTGAGCTATGGTCCTCTGCACCTTAGCCCTCG

SEQ.ID.NO. 483 HG1015084N0 PLT0001433O10.contig

GCCTGAATAATGCTGTTGTGTCACCTGAAATCTGGTCGTGTAGAGAAGTGAGTCTGCAG
GATGCTGTTCGTGTATAATATATCACAGTGGTTGCTGGAGAGGGAACCTCCACAGAGG
GATCGCTGACTTTAAGCTGCAACATAGCTCAACGGTCAGAGGTGATGTTCCACCCAAAA
TCTAACTGTGATTACTGCAGGCTACAAACTTATTTGCCCTCATGGAATAGTGAAGAACC
AAAAATGTTCACCCCAAAATATTGAGGATTGCTAATGACACTGAAAAACTCAGGGAAAACT
CTGCCCTCCACTTCTATTGCTGTATCACAGGACATAAATCCTTCCTTACTGAAAAACAGCA
CTTGTGTGTGTCCAGAGAGGCCACCAGCAGGCCAACACAGGATCTGGAAACTGATTTT
ACCATCTTTCCACATTCTCCACCTTTTAAAGACCACAAAATTGCTCTCTCCTTTGCCTTG
CCACTAAGTAGGATTTCTGGCTCTTTGTGAAAATACTATTTAAGCAAACCCCTCTAAGCTA
CCACCTTGAGAGAGAAAACTTTTGAGTAGAGGCCTTTCCACACAAATTGGTACAGTGTA
CATTAAATGTATTTCTGCTTGTTTTGTGTTAATCTGACTTTTTGGTTTCAAGAGAGTA
TCTCAACTGAGAACCTAAAAAGGAAAAAATATTATATTTTATCTTCTACAAATAGGTATG
TATTTTATCCCTATCATACAGATGAGAAAATTGAGGCACACAGGTTTTTTTTTGTAAAGAA

AGTTGATTAATAATGACATAAAGGAAATCCCGAGAGGAGCCACACACCCCTTGGAATAAAG
CAGACTGCTCTGCAGGACCCGGAGACACTCCAAATACTGTGTGTATCCACGGCTGAG
AGACCCATAGATGGTTCACATCACAGGACTCTGTGCAGACAGCCCCCAGTACCAGCTTGA
AGCCAGTAGATTGCTGGTGGCTAGACCAAGAAAGAGAGATAACAATCACTGCAGATTGG
CTCACAGGAAGCCACATCCTCAGGAAAAGGGGGAGAGTACTACATCAAGGGAACACCCCA
TGGATCAAAAGAACCTGAACAACAGCCTTCAGCCCTAGACCTTCCCTCTGACAGAGACTA
CCCAAAAGACAGGGAACCAGAAAATCAACTCTGGTAATATGACAAAACAAGCTCTTTAA
CACCCCCCAAAAATCACACTAGTTCACCAGCAGTGGATCCAAACTAAGAAGAAAATCTCT
GATTATTTGAAAAATAATTCAAGGAGGTTAGTTACCAAGCTAATCAGGGAGGCATCAGAG
AAAGGCAAAAGCCCAATGCAAGGAAATCCTAAAAAACTATACAAGAAAGCAATGAAAGAAAT
ATTCAAGGAAATATATAGCATAAAGAAAAAACTAACCAAAATCTTCAGGAAATATTGGATAC
ACTTACAGAAAGTGCAAAATGCTCTGGAAAGTCTCAACAATAGAGTTGAACAAGTAGAAGA
AAGAAATTCAAGCTCAAAGACAAAGGTCTTCAAAATTAACCCAAATCCAAACAAAGACAAAGA
AAAAAGAATAAGAAAAATATGAACAAAGCCTCTAAGAAGTCTGGGATTATGTTAAATGACA
AAACCTAAGAATAATTGGTGTTCCTGAGGAAGAAAGAGAAAATCCAAAGTTTGAAAAATATA
TTTTAGGGAATAATCGAGGAAAACCTCCCTGGCCTTGCTAGAGACACAGACATCCAAATA
CAAGAAGCACAAAGAACAACCTGGGAAATTTCATCACAAAAAGATCATCACCTAGGCACATT
ATCATCAGGGTATCTAAAGTTAAGACGAAGGAAAGAAATCTTAAGAGCTGTGAGAAAAAGC
ACCAGGTAAACCTATAAAAGAAAAACCTATCAGATTAAACAGCAGATTTCTCAACAGAAAAACC
TACAGGCTAGAAGGAAATTGGGGCCCTATCTTTAGCCTCCTCAAAACAATAATATCAGC
CAAGAAATTTGTATCCAGCGAAACTAAGCATCATATATGAAGGAAAGATACAGTTGTTTT
TCAGACTAACAAATGAGAGATTTTGCCAGTACCGAGCCACCAAGTACAAGAAATGCTAAAA

GGAGCTCTAAATCTTGAAACAAATCCTGGAAACATATCAAAATAAAACCTCTTTAAAGCA
TAAATCACACAGGACCTATAAACA

SEQ.ID.NO. 484 HG1015085N0 PLT00014333O17.contig
GGCAACAAAACCCGGAGCTGAGCGTGAAAGGCCTGGGCGAGCGGAAGGTGGGAGAGAGG
GGCTAGGCCGGGAGGGAAGCGCTGGGGCCGGGGCCGGCTTTCCAGGAGCCGGCGGAGC
GGGGTGCGTAAGAGGGGGCGAGGAAGAACCCGGCGGGCAACGGGAAGGGCAGAGGTGGAT
CTCTCCCTCAGCGACCCCTCTCGCCGAGTGAAGGTGAGGGTTTGGGGCACAAGGGCG
CAGAACGTTTTTCTGTGAGAGGAACATTAAAGACTCCTGGGAGGGCTGGGGCTAACGCC
TCCGAGAGCCCCACCTGGTCAGCCTAGGGTTAGACGTGTAGCCAGGGTGTCTGGGAGGGGG
CCGCTGCTGGCCCCAGATAGAGGGTAGAGGCCGCTCAATCACGGAGCCAGGCCCCACC
TGCTGTAGAAACCGCCACCGAGAGCCGGCCGGCTTGCTGCACGCTCTGGATTCCGCCTG
CCTATAGGCGAGGCACAAGGACTGCGAGACCAGGCCTGAAGGCAGAGAGTCAAGCAGGC
TCGGGGCCCTCTGTCCAGCGGACAACAAGCGCACCGGGCGGGAACGGCCCGCTCCCCAG
GCGCTGCAGCCTAGCTCTCATGTGACCCACGGAACCTAGCCCCCTCTGTTCTGTTGCATA
CCTGCAAAATTGTGGGCATTACGCCGGGACTGCCAACTCACCCAGGTTATTTATAGTGACA
AGAGAGTGAGACAGTGGCTTGAAAGTGTTTCCAGACCAAGGCTCTGTCCCACATATAG
GTGGATTAACTGTCTGGCAAAGGCAGGTCTCTTCTTCGGCCCTTCTGACATGGTGCCAGC
CCTACTGGAGTCAGCAGGAGTCCAGCACAAACAGATGGTGGTCAGTGGTGACTTCGGCACA
GGCAGTTCGTGAGGAGTGGTTGGTGCCCATGGAGTCCTGACTGAGTAGGTTAAAAAGAA
GGAGGAAAATCACAAGAAAGAGAGACTGGAGGTGGAGGTTGGGGAGAAAAGAGATAATGG
TATGGTGAGATGTCTACTGAACCTGCTCCTGCCAAGGCCCTCCCCATGTACTCTCAACGA

AAGTTGCAGCCACCCATTCTCATTTTGTGGATATAATTCCCTGCTTTAGTTTTCTGAGCA
CAAGAAATCAACTTTGGCTGCACATTGGAATATCTGAGAAAGCTTAAATACTGATGTTT
GGACCAGGTGCAGTGGCTCACGCCCTGTAAATCCCAGCACTTTGGGAGGCAGAGGCCAC
CTCTCAACAACCCATGAAGTGTCTGGAGCTCCACCCTCTGCTCAAGAGGAACGGCACATG
ACCCAGGCCTATGCCAATCAGCATTTTCACATTCCTCATGGTTGGTTCAAAGCTTTGTATC
TGGGTACACCTGTTTCCTCCTGAGACATGAAGGTCCTTGAGGACATTTGCTTCCTGGAGA
GAAAATTGTCACCAAAATTAAAAATACAATCTAGAGGTAGCTTTCTGAAATCATTTTAAA
CATAAAGAGATATGACTTAAAAATTTTTCATCTGAACTGTCAATTTCCATAAATAGCTTA
ATATAGTGAAAAATTGAGAGGTTCTTGAAGCCACTAAGT